

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2004, 09:34:56 ; Search time 5463 Seconds
(without alignments)
536.882 Million cell updates/sec

Title: US-10-030-194A-4

Perfect score: 2950

Sequence: 1 MKRDHLHFGQGNHGTSTAGS.....LSWTRPLITTSANKLSAVH 572

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10030194/runat_01112004.184754.28750/app.query.fasta_1.711
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10030194@cgn 1 1 912 @runat_01112004.184754.28750
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/pct_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/pctUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No. Score Match Length DB ID Description

1	2408.5	81.6	1764	16	US-10-412-699B-215	Sequence 215, App
2	2407	81.6	1764	9	US-09-938-842A-2513	Sequence 2513, Ap
3	2407	81.6	1764	11	US-09-938-842A-2513	Sequence 2513, Ap
4	2111	71.6	1602	9	US-09-938-842A-1026	Sequence 1026, Ap
5	2111	71.6	1602	11	US-09-938-842A-1026	Sequence 1026, Ap
6	2110.5	71.5	1964	9	US-09-911-513-1	Sequence 1, Appli
7	2110.5	71.5	1964	10	US-09-911-513-1	Sequence 1, Appli
8	2106.5	71.4	1951	15	US-10-278-536-45	Sequence 45, Appl
9	2106.5	71.4	1951	15	US-10-225-066A-783	Sequence 783, App
10	2106.5	71.4	1951	16	US-10-374-780A-2289	Sequence 2289, Ap
11	2106.5	71.4	1951	16	US-10-412-699B-217	Sequence 217, App
12	1943	65.9	1643	9	US-09-911-513-3	Sequence 3, Appli
13	1943	65.9	1643	10	US-09-911-513-3	Sequence 3, Appli
14	1929	65.4	1642	9	US-09-911-513-9	Sequence 9, Appli
15	1929	65.4	1642	10	US-09-911-513-9	Sequence 9, Appli
16	1923	65.2	1642	9	US-09-911-513-5	Sequence 5, Appli
17	1923	65.2	1642	10	US-09-911-513-5	Sequence 5, Appli
18	1905	64.6	1636	9	US-09-911-513-7	Sequence 7, Appli
19	1905	64.6	1636	10	US-09-911-513-7	Sequence 7, Appli
20	1788	60.6	2174	16	US-10-424-599-8444	Sequence 8444, Ap
21	1774.5	60.2	1728	16	US-10-425-114-8187	Sequence 8187, Ap
22	1748	59.3	1418	16	US-10-425-114-14732	Sequence 14732, A
23	1748	59.3	1418	16	US-10-425-114-14732	Sequence 14732, A
24	1710	58.0	2028	17	US-10-424-599-125726	Sequence 125726,
25	1594	54.0	2385	15	US-10-310-154-201	Sequence 201, App
26	1592	54.0	2500	15	US-10-310-154-194	Sequence 194, App
27	1551.5	52.6	2302	17	US-10-655-799-20	Sequence 20, Appl
28	1456.5	49.4	1383	16	US-10-425-114-11522	Sequence 11522, A
29	1448.5	49.1	1889	16	US-10-425-114-12943	Sequence 12943, A
30	1351	45.8	1568	16	US-10-425-114-31185	Sequence 31185, A
31	1350.5	45.8	1075	16	US-10-425-114-20656	Sequence 20656, A
32	1162.5	39.4	1949	16	US-10-425-114-29590	Sequence 29590, A
33	1162.5	39.4	2019	16	US-10-424-599-77544	Sequence 77544, A
34	1142	38.7	1564	16	US-10-425-114-18679	Sequence 18679, A
35	1142	38.7	1820	16	US-10-424-599-78225	Sequence 78225, A
36	1098	37.2	1362	16	US-10-425-114-17415	Sequence 17415, A
37	1050.5	35.6	1282	16	US-10-425-114-20541	Sequence 20541, A
38	964.5	32.7	909	16	US-10-425-114-19851	Sequence 19851, A
39	940	31.9	1473	17	US-10-555-799-8	Sequence 8, Appli
40	937.5	31.8	1482	17	US-10-437-963-24052	Sequence 24052, A
41	916	31.1	1117	16	US-10-425-114-12831	Sequence 12831, A
42	900	30.5	1170	16	US-10-425-114-34558	Sequence 34558, A
43	857	29.1	522	9	US-09-770-152-995	Sequence 995, App
44	823.5	27.9	655	17	US-10-021-323-5665	Sequence 5665, Ap
45	822	27.9	567	17	US-10-021-323-3979	Sequence 3979, Ap

ALIGNMENTS

RESULT 1

US-10-412-699B-215
; Sequence 215, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: Dubell, Arnold N.

APPLICANT: Ratcliffe, Oliver
APPLICANT: Kumimoto, Rodrick
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
FILE REFERENCE: MBI-0048CIP
CURRENT APPLICATION NUMBER: US/10/412,699B
CURRENT FILING DATE: 2003-04-10

PRIOR APPLICATION NUMBER: 09/394,519
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: 09/489,376
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: 09/506,720
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 09/533,030
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 09/533,392
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 09/533,029
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 09/532,591
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 09/533,648
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 09/713,994
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: 09/819,142
PRIOR FILING DATE: 2001-03-27

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 2011

SOFTWARE: PatentIn version 3.2

SEQ ID NO 215

LENGTH: 1764

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

OTHER INFORMATION: G307

US-10-412-699B-215

Alignment Scores:
Pred. No.: 3,376-263 Length: 1764
Score: 2408.50 Matches: 490
Percent Similarity: 85.79% Conservative: 23
Best Local Similarity: 81.94% Mismatches: 48
Query Match: 81.64% Indels: 37
DB: 16 Gaps: 10

US-10-030-194a-4 (1-572) x US-10-412-699B-215 (1-1764)

QY 1 MetLysArgAspLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerIleAla 18
DB 1 ATGAAGAGAGATCATCACCATTCCAAAGGTCGATTGTCACACCGGACTTCTTCTTCT 60
QY 19 GlySerSerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGlu 38
DB 61 TCATCATCAATCTCT-----AAAGATAAGATGATGATGTAAGAAA 102
QY 39 GluGlu-----AspAspGluLeuLeuGlyValLeuGlyTyrIysVal 52
DB 103 GAAGAAGACGGTGGAGGTAAACATGACGACGAGCTTCGCTGTTTATAGGTTACAAAGTT 162
QY 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuThrMetMetGly 72
DB 163 AGGTCAATCGGAGATGCGGAGGTTCCTTTGAACCTCGAACCAATTAGACGATGATGAT 222
QY 73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92
DB 223 AATGTTCAAGAAGATGGTTTATCTCATCTCGCGACGATCTGTTCAATTATAATCCGTCG 282
QY 93 GluLeuTyrSerTyrLeuAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
DB 283 GAGCTTTATCTTGGCTTGATAATATGCTCTCTGAGCTTAAATCCTCTCTCTCCGCGG 342
QY 113 GlySerAsnAlaLeuAsnProGluIleAsnAsnAsnAsnSerPhe----PheThr 131

DB 343 AGTTCTAAACGGTTTAGATCCGGTTCTTCTTCGCCGAGATTTCGTGGTTTCCGCTTCG 402
QY 132 GlyGlyAspLeuLysAlaIleProGlyAsnAlaValCysArgArgSerAsnGlyPheAla 151
DB 403 GATTATGACCTTAAAGTCATTCCCGAAACCGGATT-----TATCAGTTTCG 450
QY 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
DB 451 ---GCGATTGATTCTTCTTCGTCGAATAATCAGAACACGCTTTGAATCATGCTCG 507
QY 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182
DB 508 AGTCCTGATTCTATGTTTACATCGACTTCGACGGGTACGAGATTGGTGGAGTCATAGGA 567
QY 183 ThrThrValThrThrValThr-----GluSerThrArgProLeu 195
DB 568 ACGACGGTCAGCACACACCGACCAACGACGCGCGCGCTGAGTCAACTCGTTCGT 627
QY 196 IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
DB 628 ATCTGTTGTTGACTCGAAGACGGTGTTCGTTAGTCCACGCGCTTATGCTTGTGCA 687
QY 216 GluAlaValGlnSerSerAsnLeuThrIleuAlaGluAlaLeuValLysGlnIleGlyPhe 235
DB 688 GAAGCAATCCAGCAGACAACATTTGACTCTAGCGAAGCTCTTGTGAAGCAATCGGATGC 747
QY 236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrTyrPheAlaGluAlaLeu 255
DB 748 TTAGCTGTCTCAAGCGGAGCTTATGAGAAAGTGGCTACTTACTTCGCCGAGACTTTA 807
QY 256 AlaArgArgIleTyrArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAsp 275
DB 808 GCTCGGCGGATCTACCGTCTCTCTCCGCCGAGAAATCAGATCGATCATTTGTTCCGAT 867
QY 276 ThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295
DB 868 ACTCTTCAGATGCACCTTTTACGAGACTTGTCTTATCTTAAATTCGCTCACTTCACGCG 927
QY 296 AsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgValHisValIleAspPheSer 315
DB 928 AACCAACGATTTCTCGAAGCTTTTGAAGGTAAAGAGAGATACACGTCATTTGATTTTCG 987
QY 316 MetAsnGlnGlyLeuGlnTrrProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGly 335
DB 988 ATGAACCAAGTCTTCAATGGCTTCGCTTATGCAAGCTCTTCGCTTCGAGAAAGGAGT 1047
QY 336 ProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeu 355
DB 1048 CCTCCAACCTTTCCGGTTAAACCGGAATTTGGTCCACCGCGCGGATAATTTCTGATCATCTT 1107
QY 356 HisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyr 375
DB 1108 CATGAAGTTGTTGTGAATTAGCTCAGCTTCGGAGGCGATTACAGTAGAATTCGAATAC 1167
QY 376 ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395
DB 1168 CGTGGATTCGTTGTTTACAGCTTAGCCGATCTCGATGCTTCGATGCTTGGAGCTTAGACG 1227
QY 396 SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArg 415
DB 1228 AGCGATACGGAAGCTGTTGCGGTGAACCTCTGTTTTCAGCTTACATAAGCTCTTAGAGTCT 1287
QY 416 ThrGlyGlyIleGluLysValPheGlyValValLysGlnIleLysProValIlePheThr 435
DB 1288 CCCGGTGGGATAGAGAAAGTTCTCGCGGCTTGTGAACAGATTAACCCGGTATTTCACG 1347
QY 436 ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455
DB 1348 GTGGTTGAGCAAGAAATCGAACCATTAACGACCGGTTTTCTTAGACCGGTTTACTGAATCG 1407
QY 456 LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys 475

Db 1408 TTACATTATTATTCGACTCTGTTTGAATTCGTTGGAGAGGTTCCGAATAGTCAAGACAAA 1467
Qy 476 ValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro 495
Db 1468 GTCATGTCTGAAGTTTACTTAGGGAACACAGATTTGTAAATCTGGTGGCTTGTGAAGGTCT 1527
Qy 496 AspArgValGluArgHisGluThrLeuSerGlnTyrSerAsnArgPheGlySerSerGly 515
Db 1528 GACAGAGTCGAGAGACAGAAACGTTGAGTCAATGGGGAACCGGTTTGGTTCTGTCGCGGT 1587
Qy 516 PheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeu 535
Db 1588 TTAGCGCCGGCACATCTGGGTCTAACGGTTTAAAGCGAGTATGCTTTTGTCTGTG 1647
Qy 536 PheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGlyCysLeuMetLeuSerTrp 555
Db 1648 TTTAATAGTGGCAAGTTATCGTGTGGAGGAGAGTAAATGGATGTTTGAATGTTGGTTGG 1707
Qy 556 HisThrArgProLeuIleThrThrSerAlaTyrLysLeuSer---AlaValHis 572
Db 1708 CACACTCGCCCACTCATACCACTTCGCTTGGAAACTCTCGACGGCGCGGCAC 1761

RESULT 2
US-09-938-842A-2513
; Sequence 2513, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Kreps, Jeff
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2513
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2513

Alignment Scores:
Pred. No.: 4,99e-263 Length: 1764
Score: 2407.00 Matches: 488
Percent Similarity: 86.03% Conservative: 23
Best Local Similarity: 82.15% Mismatches: 47
Query Match: 81.53% Indels: 36
DB: 9 Gaps: 9

US-10-030-194A-4 (1-572) x US-09-938-842A-2513 (1-1764)

Qy 1 MetLysArgAspLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerIleAla 18
Db 1 ATGAAGAGAGATCATCAACCAATTCGAAGTCCAGTCCCAACCCACGGGACTTCTTCTTCA 60
Qy 19 GlySerThrThrSerProAlaValPheGlyLysAspLysMetMetValLysGlu 38
Db 61 TCATCATCAATCTCT-----AAAGATAGATGATGATGGTGAAGAAA 102
Qy 39 GluGlu-----AspAspGluLeuLeuGlyValLeuGlyTyrLysVal 52
Db 103 GAAGAAGACGGTGGAGGTAAACATGACGACGAGCTTCGCTGTTTAGTTTACAAAGTT 162
Qy 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuThrMetMetGly 72

Db 163 AGGTCAATCGGAGATCGCGAGGTTGCTTTGAAACTCGAACAAATTAGACACGATGATGAGT 222
Qy 73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisPyrAsnProAla 92
Db 223 AATGTTCAAGAAGATGGTTTATCTCATCTCGACGAGTACTGTTCATTATATATCCGTGC 282
Qy 93 GluLeuTyrSerTrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
Db 283 GAGCTTTATCTTGGCTTGAATAATATGCTCTGTAGCTTAATCTCTCTCTCTCTCTCC 342
Qy 113 GlySerAsnAlaLeuAsnProGluIleAsnAsnAsnAsnAsnSerPhe---PheThr 131
Db 343 AGTTCTAACGGTTTAGATCCGGTCTCTCTCGCGGAGATTTGGTGTTCCTCGGCTTCG 402
Qy 132 GlyGlyAspLeuLysAlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAla 151
Db 403 GATTATGACCTTAAAGTCAATTCCTCCGGAACCGGAT-----TATCAGTTTCCG 450
Qy 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
Db 451 ---GCGATTGATTCTCTGCTCTGCTGAATAATCAGAACCAAGCGTTTGAATCATGCTCG 507
Qy 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182
Db 508 AGTCTGATTTCTATGTTTACATCGACTTCGACGGTACGCAGATTTGGTGGAGTCAATAGA 567
Qy 183 ThrThrValThrThrValThr-----GluSerThrArgProLeu 195
Db 568 ACGAGCGTGACGACCAACCAACGACGACGACGCGCGGGGTGAGTCAACTGTTCTGTT 627
Qy 196 IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
Db 628 ATCTCTGGTTGATCTCGAAGAGAACGTTGCTTTAGTCCACGCGCTTATGCTTGTGCA 687
Qy 216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235
Db 688 GAAGCAATCCAGCAGAACAAATTTGACTCTAGCGGAGGCTCTTGTGAGCAATCCGATGC 747
Qy 236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrTyrPheAlaGluAlaLeu 255
Db 748 TTAGCTGTGTCTCAAGCGGAGCTATGAGAAAGTGGCTACTTACTTCTCGCCGAAGCTTTA 807
Qy 256 AlaArgIleTyrArgLeuSerProProGlnThrGlnIleAspHisSerLeuSerAsp 275
Db 808 GCGGCGGAGTCTACCGTCTCTCTCGCGCGCAGAAATCAGATCGATCATTTGCTCTCCGAT 867
Qy 276 ThrLeuGlnMetHisPheThrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295
Db 868 ACTCTTCAGATGCACCTTTTACGAGACTTGTCTTAACTTAAATTCGCTCCTCCTCAGGCG 927
Qy 296 AsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgValHisValIleAspPheSer 315
Db 928 AACCAAGCGATTCTCGAAGCTTTTGAAGGTAAGAGAGAGTACACGTCATTGATTTCTCG 987
Qy 316 MetAsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGly 335
Db 988 ATGAACCAAGGTCTTCAATGGCTTATGCAAGCTTTCGCGCTTCGAGAGAGAGGT 1047
Qy 336 ProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeu 355
Db 1048 CTCTCAACTTTCCGGTTTAAACCGGAATGCTCACCGCGCGCGGATAAATCTCATCTCT 1107
Qy 356 HisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyr 375
Db 1108 CATGAAGTTGTTGTAATTAGCTCAGCTTCGGGAGGCGATTTCAGTAGAATTCGAATAC 1167
Qy 376 ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395
Db 1168 CGTGGATTCTGTTGCTTAAACAGCTTAGCCGATCTCGATGCTTCGATGCTTTCGAGCTTAGACCG 1227
Qy 396 SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArg 415
Db 1228 ACCGATACGGAAGCTGTTGCGGTGAACCTGCTGTTTTTGGAGCTACATAAGCTCTTAGTCGT 1287

Qy 416 ThrGlyGlyIleGluLysValPheGlyValVallysGlnIleLysProValIlePheThr 435
Db 1288 CCCGGTGGGATAGAAAGTTCTCGCGCTTGTGAAACAGATTAACCGGTGATTTTCACG 1347
Qy 436 ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455
Db 1348 GTGGTTGAGCAAGAAATCAACCAATGAAACCGGCTTTCTTAGACCGGTTTACTGAAATCG 1407
Qy 456 LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys 475
Db 1408 TTACATTTATTTCGACTCTGTGTTGATTGCTTGGAGGAGGATTCGAATAGTCAAGACAAA 1467
Qy 476 ValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro 495
Db 1468 GTCAATGCTGAAGTTTACTTAGGNAACAGATTTGTATCTGTGGCTTGTGAAGTCTCT 1527
Qy 496 AspArgValGluArgHisGluThrLeuSerGlnTrpSerAsnArgPheGlySerSerGly 515
Db 1528 GACAGAGTCGAGACACGAAACGTTGAGTCAATGGGGAAACCGGTTTGGTTCGTCGGT 1587
Qy 516 PheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeu 535
Db 1588 TTAGCGCGCGCACATCTTTGGGCTTAACCGCTTAAAGCAAGCGAGTATGCTTTTGTCTGTG 1647
Qy 536 PheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGlyCysLeuMetLeuSerTrp 555
Db 1648 TTTAATAGTGGCCAAAGATTTCGTGTGAGAGAGATATGATGTTTGATGTTGGTTGG 1707
Qy 556 HisThrArgProLeuIleThrThrSerAlaTrpLysLeuSer 569
Db 1708 CACACTCGTCCACTCATTTACCACCTCCGCTTGGAACCTCTCG 1749

RESULT 3

US-09-938-842A-2513
; Sequence 2513, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2513
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2513

Alignment Scores:
Pred. No.: 4,99e-263 Length: 1764
Score: 2407.00 Matches: 488
Percent Similarity: 86.03% Conservative: 23
Best Local Similarity: 82.15% Mismatches: 47
Query Match: 81.59% Indels: 36
DB: 11 Gaps: 9

US-10-030-194A-4 (1-572) x US-09-938-842A-2513 (1-1764)

Qy 1 MetLysArgAspLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerIleAla 18
Db 1 ATGAAGAGAGATCATCACCAATTCCAAAGGTGATTTGCCAACACACGGGACTTCTTCTTCA 60

Qy 19 GlySerSerThrSerSerProAlaValPheGlyLysAspLysMetMetMetValLysGlu 38
Db 61 TCATCATCAATCTCT-----AAAGATAAGATGATGATGGTGAAAAA 102
Qy 39 GluGlu-----AspAspGluLeuLeuGlyValIleGlyTyrLysVal 52
Db 103 GAAGAAGACGGTGGAGGTAAACATGGACGAGCTTCTCGCTGTTTGTAGGTTACAAAGTT 162
Qy 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGly 72
Db 163 AGGTCATCGGAGATGGCGAGGTTCCTTGAACCTCGAACATTAGAGACGATGATGAGT 222
Qy 73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92
Db 223 AATGTTCAAGAAGATGTTTATCTCATCTCGACGAGTACTGTTCTATTATATAATCCGCTCG 282
Qy 93 GluLeuTyrSerTrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
Db 283 GAGCTTTATTTCTTGGCTTGATAATATGCTCTCTGAGCTTAATCTCTCTCTCTCTCTCT 342
Qy 113 GlySerAsnAlaLeuAsnProGluIleAsnAsnAsnAsnAsnSerPhe---PheThr 131
Db 343 AGTTCTAACGGTTAGATCCGGTTCCTTCGCCGGAGATTTGTGGTTTCCGGCTTCG 402
Qy 132 GlyGlyAspLeuLysAlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAla 151
Db 403 GATTATGACCTTAAAGTCATTCCTCGGAAACCGGATT-----TATCAGTTTCCG 450
Qy 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
Db 451 ---CGCATTTGATTTCTGCTCTCGTGAATAATCAGAACCAAGCGTTTGAATCATGTCG 507
Qy 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182
Db 508 AGTCCTGATTTATGTTTACATCGACTTCACCGGTAGCGAGATTGGTGGAGTCTATAGGA 567
Qy 183 ThrThrValThrThrValThr-----GluSerThrArgProLeu 195
Db 568 ACGACGGTGACGACACCAACGACGACGCGCGGGGTGAGTCAACTCGTCTCTGTT 627
Qy 196 IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
Db 628 ATCCTGGTTGACTCCCAAGACGCGTGTGTTTAGTCCACGCGCTTATGGCTGTGCA 687
Qy 216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235
Db 688 GAAGCAATCCAGCAGAACAAATTTGACTCTACCGAAGCTCTTGAAAGCAAAATCGGATGC 747
Qy 236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrTyrPheAlaGluAlaLeu 255
Db 748 TTAGCTGTGCTCAAGCGGAGCTATGAGAAAGTGGCTACTTACTTCGCCGAAAGCTTGA 807
Qy 256 AlaArgArgIleTyrArgLeuSerProProGlnThrGlnIleAspHisSerLeuSerAsp 275
Db 808 CGCGGGGGATCTACCGCTCTCTCTCCGCGCAGAAATCAGATCGATCTGTTCTCTCCGAT 867
Qy 276 ThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295
Db 868 ACTCTTCAGATGCACCTTTTACGAGACTTGCTTATCTTAATTCGCTCCTTCTCAGCGCG 927
Qy 296 AsnGlnAlaIleLeuGluAlaPheGluGlyLysArgValHisValIleAspPheSer 315
Db 928 AACCAAGCATTTCTCGAAGCTTTTGAAGTTAAGAGAGATACAGCTCATTTGATTTCTCG 987
Qy 316 MetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGly 335
Db 988 ATGAACCAAGGTCTTCAATGGCTGCACTTATGCAAGCTCTTGGCTTTCGAGAGAGAGGT 1047
Qy 336 ProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeu 355
Db 1048 CCTCAACTTTCCGGTTAAACCGGAATTGGTCCACCGCGCGCGGATTAATCTTGATCATCTT 1107


```
|||||
802 AAGAAAGAGTTCAATGATTCATGATTCATGAGTCAAGGCTTCAATGGCCGCGCTT 861
QY MetGlnAlaLeuAlaLeuArgGluGlyProProSerPheArgLeuThrGlyIleGly 345
Db ATGACAGGCTTTGGCGCTTCGACCTGGTGGTCTCTCTCGGTTAAACCGGAATTGGT 921
QY ProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysValLeuAlaGlnLeu 365
Db CACCGGACCGGATAATTCGATTATCTTCATGAAGTTGGGTGAAGCTGCTCATTTA 981
QY AlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAsp 385
Db GCTGAGGCGAATTCACGTTGAGTTGAGTACAGAGATTGTGGCTAACACATTAGCTGAT 1041
QY LeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSer 405
Db CTTTGATGCTTCGATGCTTGGAGCTTAGACCAAGTGAGATTGATCTGTTGGCGTTAACTCT 1101
QY ValPheGluLeuHisValLeuLeuGlyArgThrGlyGlyIleGluValValPheGlyVal 425
Db GTTTCGAGCTTCAACAGCTCTTGGGACGACTGTGGGATCGATGAAGTCTTGGTGTG 1161
QY ValLysGlnIleLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGly 445
Db GTGAATCAGATTAAACCGGAGATTTTCACTGTGGTTGACGAGGATCGAACCAATAAGT 1221
QY ProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSer 465
Db CCGATTTTCTTAGATCGGTTTACTGAGTCGTTGCAATTATTACTCGACGTTGTTTACCTCG 1281
QY LeuGluGlyAlaProSerSerClnAspLysValMetSerGluValTyrLeuGlyLysGln 485
Db TTGGAAGGTGTACCGAGTGGTCAAGACAGGTCATGTCGGAGTTTACTTGGGTAAACAG 1341
QY IleCysAsnLeuValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSer 505
Db ATCTGCAACGTTGTGGCTTGTGATGACCTGACCGAGTTGAGCGTCATGAAACGTTGAGT 1401
QY GlnTrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAla 525
Db CAGTCGGAGAACCGGTTCCGGGTCCTCGGGTTTCGGGCTGCGACATATTTGGTTTCGAATCGC 1461
QY PheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGlu 545
Db TTAAAGCAAGCGAGATGCTTTTGGCTCTGTTTCAACGGCGGTGAGGGTTATCGGGTGAG 1521
QY LysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAla 565
Db GAGAGTGACGGCTGCTCATGTTGGGTTGGCACACACACGCGCTCATAGCCACCTCGGCT 1581
QY TrpLysLeuSer 569
Db TGGAAACTCTCC 1593
```

RESULT 5

```
US-09-938-842A-1026
; Sequence 1026, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
```

```
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1026
; LENGTH: 1602
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1026

Alignment Scores:
Pred. No.: 2,01e-229 Length: 1602
Score: 2111.00 Matches: 428
Percent Similarity: 79.28% Conservative: 35
Best Local Similarity: 73.29% Mismatches: 53
Query Match: 71.56% Indels: 68
DB: 11 Gaps: 8

US-10-030-194A-4 (1-572) x US-09-938-842A-1026 (1-1602)

QY 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 1 ATGAAGAGAGATCATCATCATCATCATCAAGAT----- 36
QY 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetMetValLysGluGlu 40
Db 37 -----AAGAAGACTATGATGATGAATGAAGAAGAC 66
QY 41 Asp-----AspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGlu 56
Db 67 GACGGTAAACGGCATGGATGAGCTTCTAGCTGTTCTTGGTTTACAAGGTTAGGTCATCCGAA 126
QY 57 MetAlaGluValAlaLeuLysLeuGlnLeuGluThrMetMetGlyAsnAlaGlnGlu 76
Db 127 ATGCTGATGTTGCTCAGAAACTCGAGCAGGTTGAGTTATGATGCTTAATGTTCAAGAA 186
QY 77 AspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSer 96
Db 187 GACGATCTTTCTCAACTCGCTACTGAGACTGTTCTACTATAATCGCGGAGCTTTACACG 246
QY 97 TrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAla 116
Db 247 TGGCTTGATTTCTATGCTCACCAGCTTTAATCCTCGTCTCT----- 288
QY 117 LeuAsnProGluIleAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLys 136
Db 289 ---AACGCCGAGTAC-----GATCTTAA 309
QY 137 AlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSer 156
Db 310 GCTATTCGGGTGACGCGATTCCTC-----AATCAGTTTCGCTATCGATTTCGGCTTCT 360
QY 157 SerSer-----AsnLysArgLeuLysProSer 165
Db 361 TCGTCTAAACCAAGCGCGGAGAGATACGTATACTACAAACAAGCGGTTGAAATGCTCA 420
QY 166 SerSerProAspSerMetValThrSerProSerProAlaGlyValIleGlyThrThrVal 185
Db 421 AAC-----GGCGTCGTGGAAACCACT--- 441
QY 186 ThrThrValThrGluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyVal 205
Db 442 ACAGCGACGGCTGAGTCAACTCGGCATGTTGCTCCTGTTGACTCGCAGGAGAACGGTGTG 501
QY 206 ArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeu 225
Db 502 CGTCTCGTTCACGGCGCTTTGGCTTGGCTGCGCTGTCAGAGCTGTTCAGAAAGAGAATCTGACTGA 561
QY 226 AlaGluAlaLeuValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArg 245
Db 562 GCGGAGGCTCTGGTGAACCAATCGGATTCCTAGCCGTTTCTCAATCGGAGCGATGAGA 621
QY 246 LysValAlaThrTyrPheAlaGluAlaLeuAlaArgIleTyrArgLeuSerProPro 265
Db -----
```

Db 622 AAAGTCGCTACTTACTTCGCCGAGAGCTCTCGCGCGCGGATTTACCGTCTCTCTCGTGG 681
Qy 266 GlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCys 285
Db 682 CAGAGTCCAATCGACCACTCTCTCCGATACCTCTTCAGATGCACCTTCTACGAGACTTGT 741
Qy 286 ProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGly 305
Db 742 CCTTATCTCAAGTTCGCTCACTTCACTCGCGCAATCAAGCGATTTCTCGAAGCTTTTCAAGGG 801
Qy 306 LysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTyrProAlaLeu 325
Db 802 AGAAAGAGTTCATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 861
Qy 326 MetGlnAlaLeuAlaLeuArgGluGlyGlyProSerPheArgLeuThrGlyLeGly 345
Db 862 ATGAGGCTCTTCGCGCTTCACCTGGTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 921
Qy 346 ProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeu 365
Db 922 CCACCGGCAACCGGATTAATTCGATTAATTCGATTAATTCGATTAATTCGATTAATTCGAT 981
Qy 366 AlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAlaAsnSerLeuAlaAsp 385
Db 982 GCTGAGCGGATTCACCTGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1041
Qy 386 LeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAlaValAlaValaAsnSer 405
Db 1042 CTTGATGCTTCGATGCTTCGAGCTTAGACCAAGTAGATTAATTCGTTGCGGTAACTCT 1101
Qy 406 ValPheGluLeuHisLysLeuGlyArgThrGlyGlyIleGluLysValPheGlyVal 425
Db 1102 GTTTTCGAGCTTCACAAAGCTCTTGGACGACCTGGTGGCATCGATAAGGTTCTTGGTGTG 1161
Qy 426 ValLysGlnIleValProValIlePheThrValValGluGlnGluSerAsnHisAsnGly 445
Db 1162 GTGAATCAGATTAACCGGAGATTTTCACTGTGTGGTGGACAGATTCGAACCATTAATAGT 1221
Qy 446 ProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSer 465
Db 1222 CCGATTTCTTAGATCGGTTTACTGAGTCGTTGCATTAATTAATTAATTAATTAATTAAT 1281
Qy 466 LeuGluGlyAlaProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGln 485
Db 1282 TTGGAAGGTGTACCGAGTGGTCAAGACAAAGTCAATGTCGAGGTTTACTTGGTGAACAG 1341
Qy 486 IleCysAsnLeuValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSer 505
Db 1342 ATCTGCAAGTTGTGGCTTGTGTGAGCTTGACCTGACCGAGTTGACGATGAACGTTGAGT 1401
Qy 506 GlnTyrSerAsnArgPheGlySerGlyPheAlaProAlaHisLeuGlySerAsnAla 525
Db 1402 CAGTGGAGAACCGGTTCCGGTCTGCTGGGTTTGGCGTGCACATATTTGTTTGAATGG 1461
Qy 526 PheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGlu 545
Db 1462 TTAAAGCAAGCAGATGCTTTTGGCTCTGTTCAACGCGCGGTATCGGTTGAG 1521
Qy 546 LysAsnAsnGlyCysLeuMetLeuSerTyrHisThrArgProLeuIleThrThrSerAla 565
Db 1522 GAGAGTGGCGGCTGCTCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1581
Qy 566 TrpLysLeuSer 569
Db 1582 TGGAAACTCTCC 1593

RESULT 6

US-09-911-513-1
; Sequence 1, Application US/09911513
; Patent No. US20020049995A1
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong

; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-157
; CURRENT APPLICATION NUMBER: US/09/911,513
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-513-1

Alignment Scores:

Pred. No.: 3,18e-229 Length: 1964
Score: 2110.50 Matches: 429
Percent Similarity: 79.45% Conservative: 35
Best Local Similarity: 73.46% Mismatches: 51
Query Match: 71.54% Indels: 69
DB: 9 Gaps: 9

US-10-030-194A-4 (1-572) x US-09-911-513-1 (1-1964)

Qy 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 209 ATGAAGAGAGATCATCATCAT-----CATCATCAAGATAAG----- 244
Qy 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetMetValLysGluGlu 40
Db 245 -----AAGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 271
Qy 41 Asp-----AspGluLeuLeuGlyValLeuGlyTyrLysValArgSerGlu 56
Db 272 GACGGTAAACGGCATGGATGAGCTTCTAGCTGTTCTTGTACAGGTTAGGTCATCGGAA 331
Qy 57 MetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGlu 76
Db 332 ATGGCTGATGTTGCTCAGAAACTCGAGCAGCTTGAAGTTATGATGTTCTAAATGTTCAAGAA 391
Qy 77 AspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSer 96
Db 392 GACGATCTTTCTCAACTCGCTACTGAGACTGTTTCACTATAATATCCGCGGAGCTTTACACG 451
Qy 97 TrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaAlaThrThrGlySerAsnAla 116
Db 452 TGGCTTGTATCTATGCTCACCAGCTTAATCTCTCGCTGCTCT----- 493
Qy 117 LeuAsnProGluIleAsnAsnAsnAsnAsnSerPhePheThrGlyClyAspLeuLys 136
Db 494 ---AACGCCGAGTAC-----GATCTTAA 514
Qy 137 AlaIleProGlyAsnAlaValCysArgSerAsnGlnPheAlaPheAlaValaAspSer 156
Db 515 GCTATTCGCGGTGACGCGATTTCTC-----AATCAGTTTCGCTATCGATTCGGGTTCT 565
Qy 157 SerSer-----AsnLysArgLeuLysProSer 165
Db 566 TCGTCTTAACCAAGCGCGGAGAGATACGTATATACTACAACAAAGCGTTGAATGCTCA 625
Qy 166 SerSerProAspSerMetValThrSerProSerProAlaGlyValIleGlyThrVal 185
Db 626 AAC-----GGCGTCTGTGAAACACCACC----- 646
Qy 186 ThrThrValThrGluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyVal 205
Db 647 ACAGCGACGGTGTAGTCAACTCGGCTATGTTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 706

Qy	206	ArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeu	225
Db	707	CGTCTCGTTACGCGCTTTTGGCTTGGCTGGAAGCTGTTCAGAAGGAGAATCTGACTGTG	766
Qy	226	AlaGluAlaLeuValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArg	245
Db	767	GCGGAAGCTCTGGTGAAGCAAAATCGAATCTTACCTGTTTCTCAAAATCGGAGCTATGAGA	826
Qy	246	LysValAlaThrTyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProPro	265
Db	827	AAAGTCGTACTTACTTTCGCGAAGCTCTCGCGCGCGGATTTACCGTCTCTCTCCGTCG	886
Qy	266	GlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCys	285
Db	887	CAGAGTCCAATCGACCACCTCTCTCTCGATACTCTTCAGATGCACCTTCTACGAGACTTGT	946
Qy	286	ProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGly	305
Db	947	CGTTATCTCAAGTTTCGCTCACTTCACGCGCGAATCAAGCGATTCGCAAGCTTTTCAAGG	1006
Qy	306	LysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeu	325
Db	1007	AAGAAAGAGTTTCATGTCTATGTATTTCTCTATGAGTCAAGGCTTCTCAATGGCCGCGCTT	1066
Qy	326	MetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGly	345
Db	1067	ATGCAGGCTCTTGGGCTTCGACCTGGTGGTCTCTCTCTGTTTCCGGTTAAACCGGAATTTGGT	1126
Qy	346	ProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeu	365
Db	1127	CCACCGGACCGGATAAATTCGATATATCTTCATGAAGTTGGTGTGAAGCTGGCTCATTTA	1186
Qy	366	AlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAsp	385
Db	1187	GCTCAGCGGATTCACGTTGAGTTTGAGTACAGAGGATTTGTGGCTAACACTTTTAGCTGAT	1246
Qy	386	LeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSer	405
Db	1247	CTTGATGCTTCGATGCTTTGAGCTTTAGACCAGTGAGATGGAATCTGTGTGGGTTAACTCT	1306
Qy	406	ValPheGluLeuHisLysLeuLeuGlyValArgThrGlyGlyIleGluLysValPheGlyVal	425
Db	1307	GTTTTCAGAGCTTCAAGAGCTCTTGGGACGACCTGGTCGATCGGTAAGGTTCTTGGHGTG	1366
Qy	426	ValLysGlnIleLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGly	445
Db	1367	GTGAATCAGATTTAAACCGGAGATTTTCACTGTGCTTGACGAGGAATCGAACCAATAAGT	1426
Qy	446	ProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAsnSer	465
Db	1427	CCGATTTTCTTAGATTCGGTTTACTTGAGTCGTTGCAATATTACTCGACGTTGTTTGACTCG	1486
Qy	466	LeuGluGlyAlaProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGln	485
Db	1487	TTGAAGGTGTACCGAGTGTCAAGACAGAGTCAATGTCGAGGTTTACTTGGGTAAACAG	1546
Qy	486	IleCysAsnLeuValAlaCysGluCylProAspArgValGluArgHisGluThrLeuSer	505
Db	1547	ATCTGCAACGTTGTGGCTTGTGATGGACCTGACCGAGTTGAGCGTCATGAAACGTTGAGT	1606
Qy	506	GlnTrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHisLeuGlyLysAsnAla	525
Db	1607	CAGTGGAGGAACCGGTTTCGGGTCGCTGGGTTTCGGGCTGCACATATTGGTTGGAATGCG	1666
Qy	526	PheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGlu	545
Db	1667	TTTAAGCAAGCAGATATGCTTTTGGCTCTGTTCACCGCGGCTGAGGGTTATCGGTCGAG	1726
Qy	546	LysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAla	565
Db	1727	GAGAGTCACGGCTGCTCATGTTGGGTTGGCACACAGACCGCTCATAGCCACCTCGGCT	1786

Qy	566	TrpLysLeuSer 569			
Db	1787	TGAAACTCTCC 1798			
RESULT 7					
US-09-911-514-1					
; Sequence 1, Application US/09911514					
; Publication No. US20030084470A1					
; GENERAL INFORMATION:					
; APPLICANT: Harberd, Nicholas P					
; APPLICANT: Peng, Jinrong					
; APPLICANT: Carol, Pierre					
; APPLICANT: Richards, Donald E					
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana					
; FILE REFERENCE: 620-158					
; CURRENT APPLICATION NUMBER: US/09/911,514					
; CURRENT FILING DATE: 2001-07-25					
; PRIOR APPLICATION NUMBER: US 09/117,853					
; PRIOR FILING DATE: 1998-08-12					
; PRIOR APPLICATION NUMBER: PCT/GB97/00390					
; PRIOR FILING DATE: 1997-02-12					
; PRIOR APPLICATION NUMBER: GB 9602796.6					
; PRIOR FILING DATE: 1996-02-12					
; NUMBER OF SEQ ID NOS: 12					
; SOFTWARE: PatentIn Ver. 2.0					
; SEQ ID NO 1					
; LENGTH: 1964					
; TYPE: DNA					
; ORGANISM: Arabidopsis thaliana					
US-09-911-514-1					
Alignment Scores:					
Pred. No.:		3.18e-229	Length:	1964	
Score:		2110.50	Matches:	429	
Percent Similarity:		79.45%	Conservative:	35	
Best Local Similarity:		73.46%	Mismatches:	51	
Query Match:		71.54%	Indels:	69	
DB:		10	Gaps:	9	
US-10-030-194A-4 (1-572) x US-09-911-514-1 (1-1964)					
Qy	1	MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20			
Db	209	ATGAAGAGAGATCATCATCAT-----CATCATCAAGATAAG-----			
Qy	21	SerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGluGluGlu 40			
Db	245	-----AAGACTATGATGATGAATGAAGAAGAC 271			
Qy	41	Asp-----AspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGlu 56			
Db	272	GACGGTAACGGCATGGATGAGCTTCTACTGTTCTTGGTTACAAGTTAGGTATGCTCGAA 331			
Qy	57	MetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGlu 76			
Db	332	ATGCTGATGTTGCTCAGAACTCGAGCAGCTTGAAGTTATGATGCTTAATGTTCAAGAA 391			
Qy	77	AspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSer 96			
Db	392	GACGATCTTCTCAACTCGCTACTGAGACTGTTCACTATAATCCGCGGAGCTTTACACG 451			
Qy	97	TrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAla 116			
Db	452	TGGCTTGATTTATGCTCACCAGCCTTAATCTCGTCTGCT-----			
Qy	117	LeuAsnProGluIleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLys 136			
Db	494	---AACGCCGAGTAC-----CATCTTAAA 514			
Qy	137	AlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSer 156			
Db	515	GCTATTCGGTGACGCGATTCTC-----AATCAGTTCGCTATCATGATTCGGGTTCT 565			

Pred. No.: 8,98e-229 Length: 1951
 Score: 2106.50 Matches: 428
 Percent Similarity: 80.55% Conservative: 40
 Best Local Similarity: 73.67% Mismatches: 55
 Query Match: 71.41% Indels: 58
 DB: 15 Gaps: 8

US-10-030-194A-4 (1-572) x US-10-225-066A-783 (1-1951)

Qy 5 LeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySerThrSerSer 24
 Db ATCCATCTCTGAABAAACCCCAACCATGAGAGATCATCATCATCA----- 224
 Qy 25 ProAlaValPheGlyLys-AspLysMetMetMetValLysGluGluGluAsp----- 41
 Db -----AGATAAGAAAGACTATGATGATGAATGAAGAAAGACGCGGTAAACGG 269
 Qy 42 ---AspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVa 60
 Db CATGATGAGCTTCCTAGCTGTTCTGGTTTACAGGTTTAGGTCATCGGAAATGGCTGATG 329
 Qy 60 lAlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAl 80
 Db TGCTCAGAACTCGAGAGCTTGAAGTTATGATGCTAATGTTCAAGAGAGCATCTTTC 389
 Qy 80 aHisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTrpLeuAspAs 100
 Db TCAACTCGCTACTGAGACTGTTCACTATAATCCGCGGAGCTTTACACGTGGCTTGATTC 449
 Qy 100 nMetLeuThrGluLeuAsnProAlaAlaThrThrGlySerAsnAlaLeuAsnProGl 120
 Db TATGCTCACCCGACCTTAATCTCTCGCTCT-----AACGCCGA 488
 Qy 120 uIleAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGl 140
 Db GTAC-----GATCTTAAAGCTATTCGCG 512
 Qy 140 yAsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerSer----- 158
 Db TGACGCGATCTC-----AATCAGTTGCTATCGATCGGCTCTTCTGTAAACCA 563
 Qy 159 -----AsnLysArgLeuLysProSerSerSerProAs 169
 Db AGCGCGGAGAGATACGTATACAAACAGCGGTGAATGCTCAAC----- 615
 Qy 169 pSerMetValThrSerProAlaGlyValIleGlyThrThrValThrValTh 189
 Db -----GGGTCGTGGAAACCACC---ACAGCGACGCG 644
 Qy 189 rGluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHi 209
 Db TGAGTCAACTCGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704
 Qy 209 sAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLe 229
 Db CGCGCTTTTGGCTGCGCTGAAGCTGTTTCAAGAGAGATCTGACTGTGGCGGAGCTCT 764
 Qy 229 uValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaTh 249
 Db GGTGAGCAATCGGATTTCTAGCTGTTTCTCAATCGGAGCTATGAGCAAGTCCGTAC 824
 Qy 249 rTyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProProGlnThrGlnI 269
 Db TTACTTCCCGAAGCTCTCGCGCGCGGATTTACCGTCTCTCTCGTCGACAGAGTCCAA 884
 Qy 269 eAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLy 289
 Db CGACCACTCTCTCCGATACCTCTTTCAGATGACCTTCTACGAGACTTCTCTCTTCTCAA 944
 Qy 289 sPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgVa 309
 Db GTTCGCTCACTTCACGCGGAATCAAGCGATTTCTCGAAGCTTTTCAAGGGAAGAAAGAGT 1004

Qy 309 lHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLe 329
 Db TCATGTCATTGATTTCTCTATGAGTCAAGGCTTTCAATGGCGGGCTTATGAGGCTCT 1064
 Qy 329 uAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAl 349
 Db TGGCTTCGACCTGGTGTCTCTCTGTTTCCGGTTAACCGAATGGTTCACCGGACCC 1124
 Qy 349 aAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaI 369
 Db GGATAATTTTCGATTATCTTCATGAAGTTGGGTGTAAGCTGCTCTATTTAGCTGAGCGAT 1184
 Qy 369 eHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSe 389
 Db TCAGTTGAGTTTGAAGTACAGAGGATTTGGCTGAACACTTTAGCTGATCTTTGATGCTTC 1244
 Qy 389 rMetLeuGluLeuArgProSerGluThrGluAlaValAlaValaValaValaValaVala 409
 Db GATGCTTGAGCTTACACCAAGTGAGTTGAATCTGTTGCGGTAACTCTGTTTTCGAGCT 1304
 Qy 409 uHisLysLeuLeuGlyArgThrGlyIleGluLysValPheGlyValValLysGlnI 429
 Db TCACAAGCTCTTGGGACGACCTGGTCGATCGATAAGGTTCTTGGTGGTGAATCAGAT 1364
 Qy 429 eLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPheLe 449
 Db TAAACCGGAGATTTTCACTGTTGGTTGAGCAGAAATCGAACAATTAATAGTCCGATTTTCTT 1424
 Qy 449 uAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAl 469
 Db AGATCGGTTTACTGAGTCGTTGCAATTTACTCGACGTTGTTGACTCGTTGGAAAGGTGT 1484
 Qy 469 aProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLe 489
 Db ACCGAGTGGTCAACAGCAAGGTCATGTCGGAGGTTTACTTGGGTAAACAGATCTGCAAGT 1544
 Qy 489 uValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAs 509
 Db TGTGGCTTGTATGGACCTGACCGAGTTGACCGTTCATGAAACGTTGAGTTCAGTGGAGAA 1604
 Qy 509 nArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAl 529
 Db CCGGTCGGGTCGCTCGGTTTGGGTCGACATATTGTTGCAATGCTTTTAAAGCAAGC 1664
 Qy 529 aSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGl 549
 Db GAGTATGCTTTTGGCTCTGTTTCAACGCGGTCGAGGTTATCGGTTGGAGGAGAGTGACGG 1724
 Qy 549 yCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAlaTrpLysLeuSe 569
 Db CTGCTCTCATGTTGGGTTGGCACACACGACCGCTCATAGCCACCTCGGCTTGGAAACTCTC 1784
 Qy 569 r 569
 Db 1785 c 1785

RESULT 10
 US-10-374-780A-2289
 ; Sequence 2289, Application US/10374780A
 ; Publication No. US20040019927A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, Bradley K
 ; APPLICANT: Riechmann, Jose Luis
 ; APPLICANT: Jiang, Cai-Zhong
 ; APPLICANT: Heard, Jacqueline E
 ; APPLICANT: Haake, Volker
 ; APPLICANT: Creelman, Robert A
 ; APPLICANT: Ratcliffe, Oliver
 ; APPLICANT: Adam, Luc J
 ; APPLICANT: Reuber, T. Lynne
 ; APPLICANT: Keddies, James
 ; APPLICANT: Broun, Pierre E
 ; APPLICANT: Pilgrim, Marsha L

; APPLICANT: Dubell III, Arnold T
 ;
 ; APPLICANT: Pineda, Omaira
 ;
 ; APPLICANT: Yu, Guo-Liang
 ;
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

Alignment Scores:		
Pred. No.:	8,986-229	Length:
Score:	2106.50	Matches:
Percent Similarity:	80.55%	Conservative:
Best Local Similarity:	73.67%	Mismatches:
Query Match:	71.41%	Indels:
DB:	16	Gaps:
		1951
		428
		40
		55
		58
		8


```
Db      885 CGACCACTCTCTCCGATACCTCTTCAGATGCACCTTCTACGAGACTTGTGCTTATCTCAA 944
Qy      289 sphenalhisphenrAlaAsnGlnAlaIleuGluAlaPheGluGlyLysArgVa 309
Db      945 GTTCGCTCACTTCACGGCGAATCAAGCCGATCTCGAAGCTTTTCAAGGGAAGAAAGAGT 1004
Qy      309 lHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLe 329
Db      1005 TCATGTCATTGATTTCTCTATGAGTCAAGGCTTCNATGGCGGCGCTTATGCGGCTCT 1064
Qy      329 uAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyLeGlyProProAlaAl 349
Db      1065 TGGCTTCGACCTGGTGGTCTCTCTGTTTCCGGTTAACCGAATTGGTCCACCGGCACC 1124
Qy      349 aAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaI 369
Db      1125 GGATAAATTCGATATTATCTTCATGAAGTTGGGTGAAGCTGGCTCATTTAGCTGAGCGGAT 1184
Qy      369 eHisValGluPheGluTrpArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaLe 389
Db      1185 TCACGTTGAGTTTGATACAGAGGATTTGTGGCTAACACTTTAGCTGATCTTGATGCTTC 1244
Qy      389 rMetLeuGluLeuArgProSerGluThrGluAlaValAlaValaAsnSerValPheGluLe 409
Db      1245 GATGCTTGAGCTTAGACCAAGTGAGATTGAATCTGTTCCGGTTAACTCTGTTTCGAGCT 1304
Qy      409 uHisLysLeuLeuGlyArgThrGlyGlyLeGluLysValPheGlyValValLysGlnIl 429
Db      1305 TCACAAGCTCTTGGGACACCTGGTGGCATGATAAGGTTCTTGGTGGTGAATCAGAT 1364
Qy      429 eLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPheLe 449
Db      1365 TARACCGAGATTTTCATCTGTGTTGACAGGAATCGAACCATTAATAGTCCGATTTTCTT 1424
Qy      449 uAspArgPheThrGluSerLeuHisTyrTrpSerThrLeuPheAspSerLeuGluGlyAl 469
Db      1425 AGATCGGTTTACTGAGTCGTTGCATTATTACTCGACGTTGTTTGACTCGTTGGAAGGTGT 1484
Qy      469 aProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLe 489
Db      1485 ACCGAGTCGTCAAGACAAGGTCATGTCGGAGGTTTACTTGGGTAAACAGATCTGCAACGT 1544
Qy      489 uValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAs 509
Db      1545 TGTGCTTGATGATGACCTCGACCGAGTTGAGCGTCATGAACGTTGATGATGAGTGGAGAA 1604
Qy      509 nArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAl 529
Db      1605 CCGGTTCCGGTCTGCTGGGTTTGGCGCTGCACATATTGGTTCGAATGCGTTTAAAGCAAGC 1664
Qy      529 aSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGl 549
Db      1665 GAGTATGCTTTTGGCTCTGTTCAACGGCGGTGAGGGTTATCGGGTGGAGAGTGAACGG 1724
Qy      549 yCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAlaTrpLysLeuSe 569
Db      1725 CTGTCATGTTGGTGTGGCACACAGCACCGCTCATAGCCACCTCGGCTTGGAACTCTC 1784
Qy      569 r 569
Db      1785 C 1785
RESULT 12
US-09-911-513-3
; Sequence 3, Application US/09911513
; Patent No. US2002004995A1
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jintong
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-157
```

```
; CURRENT APPLICATION NUMBER: US/09/911,513
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-513-3
Alignment Scores:
Pred. No.: 2,71e-210 Length: 1643
Score: 1943.00 Matches: 399
Percent Similarity: 76.95% Conservative: 35
Best Local Similarity: 70.74% Mismatches: 52
Query Match: 65.86% Indels: 78
DB: 9 Gaps: 9
US-10-030-194A-4 (1-572) x US-09-911-513-3 (1-1643)
Qy      1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db      151 ATGAAGAGAGATCATCATCAT-----CATCAT-----CATCAT----- 177
Qy      21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetMetValLysGluGluGlu 40
Db      178 -----CAAGATAAGAAGACTATGATGATGAATGAAGAA 210
Qy      41 AspAspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVal 60
Db      211 GACGAC-----GGTAAACGGCATGGATGTT 234
Qy      61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db      235 GCTCAGAAACTCGAGACGTTGAAGTTATGATGCTATGCTTAATGTTCAAGAAAGACGATCTTCT 294
Qy      81 HisLeuAlaThrAspTrpValHisTyrAsnProAlaGluLeuTyrSerTrpLeuAspAsn 100
Db      295 CAACTCGCTACTGAGACTGTTCACTATAATCCGGCGGAGCTTTACACGCTGGCTTGATCT 354
Qy      101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db      355 ATGCTCACCGACCTTAATCTCCGCTCGTCT-----AACGCCGAG 393
Qy      121 IleAsnAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
Db      394 TAC-----GATCTTAAGACTATTCCTCGGT 417
Qy      141 AsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerSer----- 158
Db      418 GACCGGATCTTC-----AATCAGTTCGCTATCGATTCCGCTTCTTCGCTTAACCAA 468
Qy      159 -----AsnLysArgLeuLysProSerSerSerSerProAsp 169
Db      469 GCGCGCGGAGGAGATACGTATATACAAACAAGCGGTTGAATAATGCTCAAAC----- 519
Qy      170 SerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValThr 189
Db      520 -----GGCGTCTGTGGAAACCAACC---ACAGCGACGGCT 549
Qy      190 GluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHis 209
Db      550 GAGTCAACTCGGCATGTTGCTCTGTTGACTCGCAGGAGAACGGTGTGCGTCTCGTTCAC 609
Qy      210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229
Db      610 CGCGTTTTGGCTTGGCTGAAGCTGTTCAGAAAGGAGAAATCTGACTGTGGCGGAAGCTCTG 669
```



```
Db 469 GCGCGCGGAGAGATACGTATACAAACACGCGGTGAAATGCTCAAAC----- 519
Qy 170 SerMetValThrSerProSerProAlaGlyValIleGlyThrValThrValThr 189
Db 520 -----GGCGTCGTGGAAACACACC--ACAGCGACGCT 549
Qy 190 GluSerThrArgProLeuIleValAspSerGlnAspAsnGlyValArgLeuValHis 209
Db 550 GAGTCAACTCGGCA-TGTGTCTCTGCTTACTCGCAGAGAAACGCTGCTGCTTCAC 608
Qy 210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229
Db 609 GCGCTTTTGGCTTGGCTTGAAGCTGTTCAGAAAGGAGATCTGACTGTGGCGAAGCTCTG 668
Qy 230 ValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr 249
Db 669 GTGAAGCAAACTCGGATTTCTAGCTGTTCTCAAATCGGAGCTATGAGAAAGTCCGCTACT 728
Qy 250 TyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProProGlnThrGlnIle 269
Db 729 TACTTGGCGAAGCTCTCGCGCGGGATTTACCGTCTCTCTCCGTCGCAGAGTCCAATC 788
Qy 270 AspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLys 289
Db 789 GACCACTCTCTCTCGATACTCTTCAGATGCACCTTCTACGAGACTTGTCCCTATCTCAAG 848
Qy 290 PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysValArgVal 309
Db 849 TTCGCTCACTTCACGGCGAATCAAGGATTTCTGAAGCTTTCAGAGGAGAGAAAGAGTT 908
Qy 310 HisValIleAspPheSerMetAsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLeu 329
Db 909 CATGTCAATGATTTCTCTATGAGTCAAGTCTTCAATGCGCGCGCTTATGCAGGCTCTT 968
Qy 330 AlaLeuArgGluGlyGlyProSerPheArgLeuThrGlyIleGlyProProAlaAla 349
Db 969 GCGCTTCGACCTGGTGGTCTCTCTCTGTTTCCGGTTAAACCGGAATTTGGTCCACCGCACCG 1028
Qy 350 AspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle 369
Db 1029 GATAATTCGATATATCTTCATGAGTTGGGTGTAAGCTGGCTCATTTAGCTGAGCGGATT 1088
Qy 370 HisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSer 389
Db 1089 CACGTTGAGTTTGAACAGAGATTTGTGGCTAACACTTTAGCTGATCTTGATCTTCG 1148
Qy 390 MetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLeu 409
Db 1149 ATGCTTGAGCTTAGACCAAGTGAGATTGAATCTGTTGCGGTTAACTCTGTTTCGAGCTT 1208
Qy 410 HisLysLeuLeuGlyArgThrGlyGlyIleGluLysValPheGlyValValLysGlnIle 429
Db 1209 CACAAGCTCTTGGGACGACCTGGTCCGATCGATAGGTTCTTGGTGTGGTGAATCAGATT 1268
Qy 430 LysProValIlePheThrValValGluGlnGlnSerAsnHisAsnGlyProValPheLeu 449
Db 1269 AAACCGGAGATTTTCACTGTGGTTCAGCAGGAATCGAACCATATAATAGTCCGATTTCCTA 1328
Qy 450 AspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAla 469
Db 1329 GATCGGTTTACTGAGTCGTGCAATTAATTAATCTGACGCTGTTGTTGACTCGTTCGAAAGGTGA 1388
Qy 470 ProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeu 489
Db 1389 CCGAGTGGTCAAGCAAGCTATGTCGAGGTTTACTTGGGTAAACAGATCTGCACAGCTT 1448
Qy 490 ValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAsn 509
Db 1449 GTGGCTTGTGATGGACCTGACCGAGTTGAGGCTCATGAAACGTTGAGTCAGTGGAGGAAC 1508
Qy 510 ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAla 529
```

```
Db 1509 CGGTTCGGTCTGCTGGTTCGGCTGCACATATTGTTGGAATGCGTTTAAGCAACGG 1568
Qy 530 SerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGly 549
Db 1569 AGTATGCTTTTGGCTCTGTTCAACGGCGGTGAGGCTTATCGGTGAGAGAGTGCACGCG 1628
Qy 550 CysLeuMetLeu 553
Db 1629 TGTCTCATGTTG 1640

RESULT 15
US-09-911-514-9
; Sequence 9, Application US/09911514
; Publication No. US20030084470A1
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,514
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-514-9

Alignment Scores:
Pred. No.: 1,06e-208 Length: 1642
Score: 1929.00 Matches: 400
Percent Similarity: 76.95% Conservative: 34
Best Local Similarity: 70.92% Mismatches: 52
Query Match: 65.39% Indels: 79
DB: 10 Gaps: 9

US-10-030-194A-4 (1-572) x US-09-911-514-9 (1-1642)
Qy 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 151 ATGAAGAGAGATCATCATCAT-----CATCAT----- 177
Qy 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetMetValLysGluGluGlu 40
Db 178 -----CAAGTAAGAAGACTATGATGATGATGAATGAAGAA 210
Qy 41 AspAspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVal 60
Db 211 GACGAC-----GGTAACGGCATGATGTT 234
Qy 61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db 235 GCTCAGAAACTCGAGCAGCTTGAAGTTATGATGCTTAATGTTCAAGAAAGACGATCTTCT 294
Qy 81 HisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTrpLeuAspAsn 100
Db 295 CAATCGCTACTGAGACTGTTCACTATANTCGGCGGAGCTTTACACGTCGCTGATTTCT 354
Qy 101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db 355 ATGCTCACCGACCTTAATCTCCGTCGTC-----AACCCGAG 393
Qy 121 IleAsnAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
```

```
Db 394 TAC-----GATCTTAAAGCTATTCCCGGT 417
Qy 141 AsnAlaValCysArgSerAsnGlnPheAlaPheAlaValAspSerSerSer----- 158
Db 418 GACGCGATTCTC-----AATCAGTTTCGTATCGATTCGGCTTCTTCGTCrTAACCAA 468
Qy 159 -----AsnLysArgLeuLysProSerSerSerProAsp 169
Db 469 GCGCGCGGAGGATACGTATACTACAACAACGCGTTGAAATGCTCAAC----- 519
Qy 170 SerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValThr 189
Db 520 -----GCGCTCGTGTGAACCAACC-----ACAGCGACGGCT 549
Qy 190 GluSerThrArgProLeuIleuValAspSerGlnAspAsnGlyValArgLeuValHis 209
Db 550 GAGTCAACTCGGCA--TGTCCTCGTGTGACTCGCAGGAGAACGGTGCCTCGTCTGCAC 608
Qy 210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229
Db 609 GCGCTTTTGGCTTGGCTGAAGCTCTCAGAAGGAGNATCTGACTGTGCGGAGCTCTG 668
Qy 230 ValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr 249
Db 669 GTGAAGCAAATCGGATTCTTAGCTGTTCTCAAAATCGGAGCTATCAGAAAAAGTCGCTACT 728
Qy 250 TyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProProGlnThrGlnIle 269
Db 729 TACTTCGCGGAAGCTCTCGCGGGCGGATTTACCGCTCTCTCCGTCGCAGAGTCCAATC 788
Qy 270 AspHisSerLeuSerAspThrLeuGlnMethHisPheTyrGluThrCysProTyrLeuLys 289
Db 789 GACCACTCTCTCCGATACTCTTCAGATGCACCTCTACGAGACTTGTCTTAICTCAAG 848
Qy 290 PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysArgVal 309
Db 849 TTCGCTCACTTCACGCGCAATCAGCGATTCTCGAAGCTTTTCAAGGGAAGAAAGATT 908
Qy 310 HisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeu 329
Db 909 CATGTCATTTGATTTCTCTATGAGTCAAGGTCTTCAATGGCGCGGCTTATCGAGGCTTT 968
Qy 330 AlaLeuArgGluGlyGlyProSerPheArgLeuThrGlyIleGlyProProAlaAla 349
Db 969 GCGCTTCGACCTGGTGGTCTCTCTCTTCCGGTTAACCGGATTTGGTCCACCGGCACCG 1028
Qy 350 AspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle 369
Db 1029 GATAATTTTCGATTATCTTCATGAAGTTGGGTGTAAAGCTGGCTCATTTAGCTGAGCGGATT 1088
Qy 370 HisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSer 389
Db 1089 CACGTTGAGTTTGAGTCAGAGGATTTGGGTAAACACTTTAGCTGATCTTGATGCTTCG 1148
Qy 390 MetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLeu 409
Db 1149 ATGCTTTGAGCTTAGACCAAGTGAGATTGAATCTGTTCGGTTAACTCTGTTTCGAGCTT 1208
Qy 410 HisLysLeuLeuGlyArgThrGlyGlyIleGluLysValPheGlyValValLysGlnIle 429
Db 1209 CACAAAGCTCTTTGGGACGACCTGGTCGATCGATAAGGTTCTTGGTGTGTAATCAGATT 1268
Qy 430 LysProValIlePheThrValValGlnGlnSerAsnHisAsnGlyProValPheLeu 449
Db 1269 AAACCGGAGATTTTCACTGTGGTTGAGCAGGAATCGAACCAATAATAGTCGGAATTTCTTA 1328
Qy 450 AspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAla 469
Db 1329 GATCGGTTTACTGAGTCGTTGCAATTATTACTCGACAGCTTGTGACTCGTTGGAAGGTGTA 1388
Qy 470 ProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeu 489
Db 1389 CCGAGTGGTCAAGACAAGGTCAATGTGCGAGGTTTACTTGGGTAAACAGATCTGCAACGTT 1448
```

```
Qy 490 ValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAsn 509
Db 1449 GTGCTTGTGATGAGACTGACCGAGTTGAGCGTCATGAAACGTTGAGTCAGTGGAGGAAC 1508
Qy 510 ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAla 529
Db 1509 CCGTTTCGGGTCCTGCTGGGTTTTCGGCTGCACATATTGGTTCGAATGCGTTTAAAGCAAGCG 1568
Qy 530 SerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGly 549
Db 1569 AGTATGCTTTTGGCTCTGTTCAACGCGCGGTGAGGTTATCGGTTGGAGGAGTACGCGC 1628
Qy 550 CysLeuMetLeu 553
Db 1629 TGTCTCATGTTG 1640
```

Search completed: November 3, 2004, 15:09:28
Job time : 5505 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2004, 04:31:51 ; Search time 160 Seconds
(without alignments)
2541.070 Million cell updates/sec

Title: US-10-030-194A-4
Perfect score: 2950
Sequence: 1 MKRDLHQFGPNHCTAGS.....LSWTRPLTTSAWKLSAVH 572

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10030194/runat_01112004.184753.28669/app_query.fasta_1.711
-DB=Issued Patents NA -QFWT=fasap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=diom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10030194 @CGN 1 1 128 @runat_01112004.184753.28669 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2110.5	71.5	1964	3	US-09-117-853-1
2	2110.5	71.5	1964	4	US-09-911-154-1
3	2110.5	71.5	1964	4	US-09-911-514-1
4	1943	65.9	1643	3	US-09-117-853-3
5	1943	65.9	1643	4	US-09-911-154-3
6	1943	65.9	1643	4	US-09-911-514-3
7	1929	65.4	1642	3	US-09-117-853-9
8	1929	65.4	1642	4	US-09-911-154-9
9	1929	65.4	1642	4	US-09-911-514-9
10	1923	65.2	1642	3	US-09-117-853-5
11	1923	65.2	1642	4	US-09-911-154-5
12	1923	65.2	1642	4	US-09-911-514-5

13	1905	64.6	1636	3	US-09-117-853-7	Sequence 7, Appli
14	1905	64.6	1636	4	US-09-911-154-7	Sequence 7, Appli
15	1905	64.6	1636	4	US-09-911-514-7	Sequence 7, Appli
16	1602	54.3	2255	4	US-09-485-529-15	Sequence 15, Appl
17	1568	53.2	2125	4	US-09-485-529-14	Sequence 14, Appl
18	1390.5	47.1	2709	4	US-09-485-529-3	Sequence 3, Appli
19	1325	44.9	1768	4	US-09-485-529-13	Sequence 13, Appl
20	1205	40.8	1746	4	US-09-485-529-57	Sequence 57, Appl
21	566	19.2	2163	3	US-09-186-276B-1	Sequence 1, Appli
22	566	19.2	2163	4	US-08-842-445-1	Sequence 1, Appli
23	566	19.2	2163	4	US-09-186-188B-1	Sequence 1, Appli
24	516.5	17.5	1094	3	US-09-186-276B-26	Sequence 26, Appl
25	516.5	17.5	1094	4	US-08-842-445-26	Sequence 26, Appl
26	516.5	17.5	1094	4	US-09-186-188B-26	Sequence 26, Appl
27	507.5	17.2	1085	3	US-09-186-276B-18	Sequence 18, Appl
28	507.5	17.2	1085	4	US-08-842-445-18	Sequence 18, Appl
29	507.5	17.2	1085	4	US-09-186-188B-18	Sequence 18, Appl
30	450.5	15.3	1368	3	US-09-186-276B-22	Sequence 22, Appl
31	450.5	15.3	1368	4	US-08-842-445-22	Sequence 22, Appl
32	450.5	15.3	1368	4	US-09-186-188B-22	Sequence 22, Appl
33	428	14.5	1093	3	US-09-186-276B-53	Sequence 53, Appl
34	428	14.5	1093	4	US-08-842-445-53	Sequence 53, Appl
35	428	14.5	1093	4	US-09-186-188B-53	Sequence 53, Appl
36	426.5	14.5	2635	3	US-09-186-276B-57	Sequence 57, Appl
37	426.5	14.5	2635	4	US-08-842-445-57	Sequence 57, Appl
38	426.5	14.5	2635	4	US-09-186-188B-57	Sequence 57, Appl
39	417	14.1	1231	3	US-09-186-276B-20	Sequence 20, Appl
40	417	14.1	1231	4	US-08-842-445-20	Sequence 20, Appl
41	417	14.1	1231	4	US-09-186-188B-20	Sequence 20, Appl
42	392.5	13.3	1928	3	US-09-186-276B-55	Sequence 55, Appl
43	392.5	13.3	1928	4	US-08-842-445-55	Sequence 55, Appl
44	392.5	13.3	1928	4	US-09-186-188B-55	Sequence 55, Appl
45	391	13.3	377	4	US-09-485-529-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1
US-09-117-853-1
; Sequence 1, Application US/09117853
; Patent No. 6307126
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1964
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-117-853-1

Alignment Scores:
Pred. No.: 1.78e-233 Length: 1964
Score: 2110.50 Matches: 429
Percent Similarity: 79.45% Conservative: 35
Best Local Similarity: 73.46% Mismatches: 51
Query Match: 71.54% Indels: 69
DB: 3 Gaps: 9

US-10-030-194A-4 (1-572) x US-09-117-853-1 (1-1964)

Qy 1 MetLysArgAspLeuHisGlnPheClnGlyProAsnHisGlyThrSerIleAlaGlySer 20

[illegible]

Alignment Scores:

Pred. No.:	1.78e-233	Length:	1964
Score:	2110.50	Matches:	429
Percent Similarity:	79.45%	Conservative:	35
Best Local Similarity:	73.46%	Mismatches:	51
Query Match:	71.54%	Indels:	69
DB:	4	Gaps:	9

US-10-030-194A-4 (1-572) x US-09-911-154-1 (1-1964)

[illegible]

RESULT 3

US-09-911-514-1
; Sequence 1, Application US/09911514
; Patent No. 6794560
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of *Arabidopsis thaliana*
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,514
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390

; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1964
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-514-1

Alignment Scores:

Pred. No.: 1,78e-233 Length: 1964
Score: 2110.50 Matches: 429
Percent Similarity: 79.45% Conservative: 35
Best Local Similarity: 73.46% Mismatches: 51
Query Match: 71.54% Indels: 69
DB: 4 Gaps: 9

US-10-030-194a-4 (1-572) x US-09-911-514-1 (1-1964)

```
QY 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
   |||||
DB 209 ATGAAGAGAGATCATCATCAT-----CATCATCAAGATAAG----- 244
QY 21 SerThrSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
   |||||
DB 245 -----AAGACTATGATGATGAATGAAGAAGAC 271
QY 41 Asp-----AspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGlu 56
   |||||
DB 272 GACGGTAACGGCATGGATGAGCTTCTAGCTGTTCTTGTTTACAAAGTTTAGGTTCATCGGAA 331
QY 57 MetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGlu 76
   |||||
DB 332 ATGGCTGATGTTGCTTCAGAAACTCGAGCAGCTTGAAGTTATGATGTTCTTAATGTTCAAGAA 391
QY 77 AspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSer 96
   |||||
DB 392 GACGATCTTCTCAACTCGCTACTGAGACTGTTCACTATAATCCGGCGAGCTTTACACG 451
QY 97 TrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAla 116
   |||||
DB 452 TGGCTTGATTTCTATGCTCACCGACCTTAATCTCTCGTCTGCT----- 493
QY 117 LeuAsnProGluIleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLys 136
   |||||
DB 494 ---AACGCCGAGTAC-----GATCTTAA 514
QY 137 AlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSer 156
   |||||
DB 515 GCTATTCCCGGTGACCGCATTTCTC-----AATCAGTTTCGCTATCGATTCGGCTTCT 565
QY 157 SerSer-----AsnLysArgLeuLysProSer 165
DB 566 TCGTCTAACCAAGGGCGGAGGAGATACGTATATACTACAAACAAGCGGTTGAAATGCTCA 625
QY 166 SerSerProAspSerMetValThrSerProSerProAlaGlyValIleGlyThrThrVal 185
   |||||
DB 626 AAC-----GGCGTCTGTGGAACCAACC----- 646
QY 186 ThrThrValThrGluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyVal 205
   |||||
DB 647 ACAGCGAGCGGTGAGTCAACTCGGCATGTTGCTTGGTTGACTCCAGAGGAACCGGTGTG 706
QY 206 ArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeu 225
   |||||
DB 707 CGTCTCGTTACGGCGCTTTTGCTTTGGCTGAGCTGTTGAAAGCTGTTTCAAGAGGAGAACTGACTGTG 766
QY 226 AlaGluAlaLeuValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArg 245
   |||||
DB 767 GCGGAAGCTCTGGTGAAGCAAAATCGGATTTCTTAGCTGTTTCTCAAAATCGGAGCTATGAGA 826
```

```
QY 246 LysValAlaThrTyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProPro 265
   |||||
DB 827 AAAGTCGCTACTTACTTTCGCCGAAGCTCTCGCGCGGGGATTTACCGTCTCTCTCCGTCG 886
QY 266 GlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCys 285
   |||||
DB 887 CAGAGTCCAATCGACCACTCTCTCTCCGATATCTCTCAGATGCACCTTCTAGGAGACTTGT 946
QY 286 ProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGly 305
   |||||
DB 947 CCTTATCTCAAGTTCGCTCACTTCACGGCGGAATCAAGCGATTTCTCGAAGCTTTTCAAGG 1006
QY 306 LysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeu 325
   |||||
DB 1007 AAGAAAGAGGTTTCATGTCATTTCTCTATGAGTCAAGGCTTCAATGCGCGCGCTT 1066
QY 326 MetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGly 345
   |||||
DB 1067 ATGAGGCTCTTGGCGCTTCGACCTGCTGGTGGTCTCTCTCTGTTTCCGGTTAACCGGAATTGGT 1126
QY 346 ProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeu 365
   |||||
DB 1127 CCACGGCAGCGGATAATTTTCGATTTATCTTCATGAAGTTGGGTGAAGCTGCTCATTTA 1186
QY 366 AlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAsp 385
   |||||
DB 1187 GCTGAGCGGATTCACGTTGAGTTGAGTACAGAGGATTTGTGGCTAACACTTTAGTCTAT 1246
QY 386 LeuAspAlaSerMetLeuLeuLeuArgProSerGluThrGluAlaValAlaValAsnSer 405
   |||||
DB 1247 CTTGATGCTTCGATGCTTTGAGCTTAGACCAAGTAGAGTTGAATCTGTTGCGGTTAACTCT 1306
QY 406 ValPheGluLeuHisLysLeuLeuGlyValArgThrGlyGlyIleGluLysValPheGlyVal 425
   |||||
DB 1307 GTTTTCGAGCTTCAAAAGCTCTTGGGAGACCTGTCGATCGATGATGAGGTTCTTGGTGTG 1366
QY 426 VallysGlnIleLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGly 445
   |||||
DB 1367 GTGAATCAGATTTAAACCGGAGATTTTCACCTGCTGGTTGAGCAGGAATCGAACATAAGT 1426
QY 446 ProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSer 465
   |||||
DB 1427 CCGATTTTCTTAGATCGGTTTACTTGAGTCGTTGCAATTTACTCGACGTTGTTTGACTCG 1486
QY 466 LeuGluGlyAlaProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGln 485
   |||||
DB 1487 TTGAAGGTGTACCGAGTGTCAAGACAAAGTCAATGTCGAGGTTTACTTGGGTAAACAG 1546
QY 486 IleCysAsnLeuValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSer 505
   |||||
DB 1547 ATCTGCAACGTTGTTGGCTTGTGATGGACCTGACCGAGTTGAGCGCTCATGAAACGTTGAGT 1606
QY 506 GlnTrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAla 525
   |||||
DB 1607 CAGTGGAGGAACCGGTTTCGGGCTCTGCGGTTTCGGCTGCACATATTTGGTTCGAATCGC 1666
QY 526 PheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGlu 545
   |||||
DB 1667 TTTAAGCAAGCGAGTATGCTTTTGCTCTGTTTCAACCGCGGTGAGGGTTATCGGGTGGAG 1726
QY 546 LysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAla 565
   |||||
DB 1727 GAGAGTGAAGCGCTGCTCATGTTGGGTTGGCACACACGACCGCTCATAGCCACCTCGGCT 1786
QY 566 TrpLysLeuSer 569
   |||||
DB 1787 TGGAAACTCTCC 1798
```

RESULT 4

US-09-117-853-3
; Sequence 3, Application US/09117853
; Patent No. 6307126
; GENERAL INFORMATION:

```

; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; CURRENT FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-117-853-3

Alignment Scores:
Pred. No.: 3,066-214 Length: 1643
Score: 1943.00 Matches: 399
Percent Similarity: 76.95% Conservative: 35
Best Local Similarity: 70.74% Mismatches: 52
Query Match: 65.86% Indels: 78
DB: 3 Gaps: 9

US-10-030-194A-4 (1-572) x US-09-117-853-3 (1-1643)
QY 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
DB 151 ATGAAGAGAGATCATCATCAT-----CATCAT----- 177
QY 21 SerThrSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
DB 178 -----CAAGATAAGAGACTATGATGATGAATGAAGAA 210
QY 41 AspAspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVal 60
DB 211 GACGAC-----GGTAACGGCATGGATGTT 234
QY 61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
DB 235 GCTCAGAAACTCGACGAGCTTGAAGTTATGATGTTCTTAATGTTCAAGAAAGACGATCTTCT 294
QY 81 HisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerThrLysAspAsn 100
DB 295 CAACTCGCTACTGACGACTGTTCACTATATATCCGGCGGAGCTTTACACGTTGGCTTGATCT 354
QY 101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
DB 355 ATGCTCACGACCTTAATCTCCGTGCT-----AACGCCGAG 393
QY 121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
DB 394 TAC-----GATCTTAAAGACTATTCGCCGT 417
QY 141 AsnAlaValCysArgSerAsnGlnPheAlaPheAlaValAspSerSerSer----- 158
DB 418 GACCGGATTC-----AATCAGTTCCGCTATCGATTCCGGCTTTTCGCTCAACAA 468
QY 159 -----AsnLysArgLeuLysProSerSerSerProAsp 169
DB 469 GCGCGCGGAGAGATACGTATACTACAAACAGCGGTTGAATGCTCAAC----- 519
QY 170 SerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValThr 189
DB 520 -----GCGCTCGTGGAAACCAACC-----ACAGCGACGGCT 549
QY 190 GluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHis 209
DB 550 GAGTCAACTCGGCATGTTGCTCTGCTGACTCGCAGGAGAACGGTGTGCTCTGTTTAC 609

```

```

210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229
DB 610 GCGCTTTTGGCTGGCTCAAGCTTTCAGAGGAGAAATCTGACTGTGCGCGAAGCTCTG 669
QY 230 ValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr 249
DB 670 GTGAAGCAAAATCGGATTTCTAGCTGTTTCTCAATCGGAGCTATGAGAAAGTCGCTACT 729
QY 250 TyrPheAlaGluAlaLeuAlaArgIleTyrArgLeuSerProGlnThrGlnIle 269
DB 730 TACITTCGCGAAGCTCTCGCGCGGAGTTTACCGTCTCTCTCCGTCGAGAGTCCAAATC 789
QY 270 AspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLys 289
DB 790 GACCACCTCTCTCGGATCTTTAGATGCACTTCTACGAGACTTGTCTTATCTCAAG 849
QY 290 PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysValArgVal 309
DB 850 TTCGCTCACTTCACGCGGAATCAACGATTTCTGAAGCTTTTCAAGGGGAGAAAGAGTT 909
QY 310 HisValIleAspPheSerMetAsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLeu 329
DB 910 CATGTCATTGATTTCTCTATGAGTCAAGTCTTCAATGCGCGGCTTATGAGGCTCTT 969
QY 330 AlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAla 349
DB 970 GCGCTTCGACCTGGTGGTCTCTCTTTTCCGGTTAAACCGGAATTTGGTCCACCGCACCG 1029
QY 350 AspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle 369
DB 1030 GATAATTTTCGATTTTCTTCATGAAGTGGGTGTAAGCTGGCTCATTTAGCTGAGCGGAT 1089
QY 370 HisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSer 389
DB 1090 CACGTTGAGTTTGAGTACAGAGGATTTGTGGCTAACACTTTAGCTGATCTTCTGCTTCG 1149
QY 390 MetLeuGluLeuArgProProSerGluThrGluAlaValAlaValAsnSerValPheGluLeu 409
DB 1150 ATGCTTGAGCTTAGACCAAGTGAATGAACTGTTGGGTTAACTCTGTTTTCGAGCTT 1209
QY 410 HisLysLeuLeuGlyArgThrGlyIleGluLysValPheGlyValValLysGlnIle 429
DB 1210 CACAGCTCTTGGGACGACCTGGTGGATCGAATAAGGTTCTTGGTGGTGGTGAATCAGATT 1269
QY 430 LysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPheLeu 449
DB 1270 AAACCGGAGATTTTACGTGGTTGAGCAGGATCGAACCATTAATAGTCCGATTTCTTA 1329
QY 450 AspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSerLeuGlyAla 469
DB 1330 GATCGGTTTACTGAGTCGTTGATTTACTCGAGCTTGTGTTGACTCGTTGGAAGGTGTA 1389
QY 470 ProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeu 489
DB 1390 CCGAGTGTCTAAGCAAGGTCATGTCGAGGTTTACTTGGGTAACAGATCTGCAACGTT 1449
QY 490 ValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTyrSerAsn 509
DB 1450 GTGGCTGTGTGAGACCTGACGAGTTGAGGCTCATGAACGTTGAGTCACTGGAGGAAC 1509
QY 510 ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAla 529
DB 1510 CGGTTCCGGTCTGCTGGGTTTTCGCGCTGCACATATGTTGTAATCGGTTTAAAGCAAGCG 1569
QY 530 SerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGly 549
DB 1570 AGTATGCTTTTGGCTCTGTTCAACGCGGTTAGGGTTATTCGGGTGGAAGAGATGACGCG 1629
QY 550 CysLeuMetLeu 553
DB 1630 TGTCTCAATGTTG 1641

```

RESULT 5

US-09-911-154-3
; Sequence 3, Application US/09911154
; Patent No. 6478809
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911.154
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-154-3

Alignment Scores:

Pred. No.:	3,06e-214	Length:	1643
Score:	1943.00	Matches:	399
Percent Similarity:	76.95%	Conservative:	35
Best Local Similarity:	70.74%	Mismatches:	52
Query Match:	65.86%	Indels:	78
DB:	4	Gaps:	9

US-10-030-194A-4 (1-572) x US-09-911-154-3 (1-1643)

Qy	1	MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer	20
Db	151	ATGAAGAGAGATCATCATCAT-----CATCAT-----	177
Qy	21	SerThrSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu	40
Db	178	-----CAAGATGAAGAGACTATGATGATGAATGAATGAAGAA	210
Qy	41	AspAspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVal	60
Db	211	GACGAC-----GGTAACGGCATGATGTT	234
Qy	61	AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla	80
Db	235	GCTCAGAAACTCGAGCAGCTTGAAGTTATGATGCTAATGTTCAAGAGAGACTTTCT	294
Qy	81	HisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTyrLeuAspAsn	100
Db	295	CAACTCGCTACTGAGACTGTTCACTATAATCCGGCGGAGCTTTACACGCTGCTGATTCT	354
Qy	101	MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu	120
Db	355	ATGCTCACCGACCTTAATCTCCCGTCT-----AACGCCGAG	393
Qy	121	IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly	140
Db	394	TAC-----GATCTTAAGCTATTCCTCCGCT	417
Qy	141	AsnAlaValCysArgSerAsnGlnPheAlaValAspSerSerSer	158
Db	418	GACGGGATTCTC-----AATCAGTTCTGCTATCGATTCCGCTTCTTCGCTAACCAA	468
Qy	159	-----AsnLysArgLeuLysProSerSerSerProAsp	169
Db	469	GGCGCGGAGGAGATACGTATACTACAAACAACGCGTTGAATGCTCAAC	519

Qy	170	SerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrThrValThr	189
Db	520	-----GGCGTCTGGAAACACACC-----ACAGCAGCGCT	549
Qy	190	GluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHis	209
Db	550	GAGTCAACTCGGCATGTTGCTCGTTGACTCGCAGGAGAACGGTGTGCTCTCGTTCAC	609
Qy	210	AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu	229
Db	610	CGCTTTTGGCTCGCTGAGCTGTTCAAGAGAGAACTCTGACTGTGCGGAGAGCTCTG	669
Qy	230	ValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr	249
Db	670	GTGAAGCAAAATCGGATTTCTTAGCTGTTTCTCAAAATCGAGCTATGAGAAAAGTCGCT	729
Qy	250	TyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProGlnThrGlnIle	269
Db	730	TACTTCGCCGAAGCTCTCGCGCGGATTTACCGTCTCTCTCGCTCGCAGAGTCCAATC	789
Qy	270	AspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLys	289
Db	790	GACCACTCTCTCCGATACCTTTAGATGCACTTCTACGAGACTTGTCTTATCTCAAG	849
Qy	290	PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysArgVal	309
Db	850	TTGCTCACCTTCACGCGCAATCAAGCATTCGAAAGCTTTTCAAGGGAAGAAAGAGTT	909
Qy	310	HisValIleAspPheSerMetAsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLeu	329
Db	910	CATGTCATTTGATTTCTCTATGAGTCAAGGTTCTCAATGGCGCGCTTATCAGGCTCT	969
Qy	330	AlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAla	349
Db	970	CGCTTCGACTGCTGCT	1029
Qy	350	AspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle	369
Db	1030	GATAATTTTCGATTTATTTTCATGAAGTTGGGTGTAAGCTGGCTCATTTAGCTGAGCG	1089
Qy	370	HisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSer	389
Db	1090	CACGTTGAGTTTGTAGTACAGAGGATTTGCGCTAACACTTTAGCTGATCTTTGATGCT	1149
Qy	390	MetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLeu	409
Db	1150	ATGCTTGAGCTTAGACCAAGTGAATGATCTGTCGGTTAACTCTGTTTTCGAGCTT	1209
Qy	410	HisLysLeuLeuGlyArgThrGlyGlyIleGluLysValPheGlyValValLysGlnIle	429
Db	1210	CACAAGCTCTTGGGACGACCTGGTGCATGATTAAGGTTCTTGGTGTGTTGAATCAGAT	1269
Qy	430	LysProValIlePheThrValGluGlnGlnSerAsnHisAsnGlyProValPheLeu	449
Db	1270	AAACCGGAGATTTTCACTGTGTTGAGCAGGAATCGAACCAATAATAGTCCGATTTCTTA	1329
Qy	450	AspArgPheThrGluSerLeuHisTyrSerThrLeuPheAspSerLeuGluGlyAla	469
Db	1330	GATCGTTTACTGAGTCTGTCATATTTACTCGAGCTTGTGTTGACTCGTTCGAGAGTCTA	1389
Qy	470	ProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeu	489
Db	1390	CCGAGTGTCTCAAGCAAGGTCATGTCGAGGTTTACTTTGGGTAAACAGATCTGCAACG	1449
Qy	490	ValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTyrSerAsn	509
Db	1450	GTGCTTGTGTGATGAGCTGACCGAGTTGAGCGTTCATGAACGTTGAGTTCAGTGGAG	1509
Qy	510	ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAla	529
Db	1510	CGTTTCGGTCTCTGCTGGGTTTCGGCTGCACATATTTGTTTGAATGCTTTAAGCAAG	1569
Qy	530	SerThrLeuLeuAlaLeuPheAsnGlyGlyGlyTyrArgValGluLysAsnAsnGly	549

```
Db 1570 AGTATGCTTTTGGCTCTGTTCAACGGCGGTGAGGTTATCGGGTGGAGGAGTGAACGGC 1629
Qy 550 CysLeuMetLeu 553
Db 1630 TGTCTCATGTTG 1641

RESULT 6
US-09-911-514-3
; Sequence 3, Application US/09911514
; Patent No. 6794560
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richard, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,514
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-514-3

Alignment Scores:
Pred. No.: 3,06e-214 Length: 1643
Score: 1943.00 Matches: 399
Percent Similarity: 76.95% Conservative: 35
Best Local Similarity: 70.74% Mismatches: 52
Query Match: 65.86% Indels: 78
DB: 4 Gaps: 9

US-10-030-194A-4 (1-572) x US-09-911-514-3 (1-1643)
Qy 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 151 ATGAAGAGAGATCATCAT- -----CATCAT----- 177
Qy 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
Db 178 -----CAAGATAAGAGAGACTATCATGATGATGAATGAAGAA 210
Qy 41 AspAspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVal 60
Db 211 GACGAC-----GGTAACGGCATGGATGTT 234
Qy 61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db 235 GCTCAGAAACTCGAGCAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGAT 294
Qy 81 HisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTrpLeuAspAsn 100
Db 295 CAACCTCGCTACTGAGACTGTTTCACTATAATCCGGCGGAGCTTTACACGCTGGCTTGAATCT 354
Qy 101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db 355 ATGCTCACCGACCTTAATCCCTCCGCTGCT-----AACGCCGAG 393
Qy 121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
Db 394 TAC-----GATCTTAAGCTATTCCCGGT 417
Qy 141 AsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSer----- 158
```

```
Db 418 GACCGGATTCTC-----AATCAGTTCGCTATTCGATTTCGGCTTCTTCGCTTAACCAA 468
Qy 159 -----AsnLysArgLeuLysProSerSerSerProAsp 169
Db 469 GCGCGCGGAGAGATACGATATACAAACAAGCGGTGAATGCTCAAC----- 519
Qy 170 SerMetValThrSerProSerProAlaGlyValIleGlyThrValThrValThr 189
Db 520 -----GCGCTCGTGAACACACC-----ACAGCGACGGCT 549
Qy 190 GluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHis 209
Db 550 GAGTCAACTCGGCATGTGCTCGTTGACTCGCAGGAGAACGGTGTGCGTCTCGTTCAC 609
Qy 210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229
Db 610 GCGCTTTGGCTTGGCTGAGCTGTTCAAGAAGAGAGATCTGACTGTGGCGGAGCTCTG 669
Qy 230 ValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr 249
Db 670 GTGAAGCAAAATCGGATCTTAGCTGTTTCTCAAATCGGAGCTATGAGAAAAGTCTACT 729
Qy 250 TyrPheAlaGluAlaLeuAlaArgIleTyrArgLeuSerProProGlnThrGlnIle 269
Db 730 TACTTCGCCGAAGCTCTCGCGCGGAGTTTACCGTCTCTCTCCGTCGACAGTCCAATC 789
Qy 270 AspHisSerLeuSerAspThrLeuGlnMethHisPheTyrGluThrCysProTyrLeuLys 289
Db 790 GACCACCTCTCTCGATCTCTTTAGATGACACTTCTACGAGACTTGTCTTCTTCTCAAG 849
Qy 290 PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysArgVal 309
Db 850 TTCGCTCACCTTCAAGCGGAAATCAAGCGATTCTCGAAGCTTTTCAAGGGAAGAAAAGATT 909
Qy 310 HisValIleAspPheSerMetAsnGlnGlyLeuGlnThrProAlaLeuMetGlnAlaLeu 329
Db 910 CATGCTCATTTGATTTCTCTATGAGTCAAGGCTTCAATGGCGCGGCTTATCGAGGCTCTT 969
Qy 330 AlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAla 349
Db 970 GCGCTTCGACCTGGTGTCTCTCTGTTTCCGGTTAACCGGAATGTGTCACCGGACCG 1029
Qy 350 AspAsnSerAspHisLeuGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle 369
Db 1030 GATAATTTTCGATTATCTTCAAGAGTTGGGTGAAGCTGGCTCATTTAGCTGAGCGGATT 1089
Qy 370 HisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspLysSer 389
Db 1090 CACGTTGAGTTTGAGTACAGAGGATTTGTGGCTAACACTTTAGCTGATCTTGTATGCTTCG 1149
Qy 390 MetLeuGluLeuArgProSerGluThrGluAlaValAlaValAlaAsnSerValPheGluLeu 409
Db 1150 ATGCTTGAGCTTAGACCAAGTGAATGATCTGTTCGGGTAACTCTGTTTTCAGGCTT 1209
Qy 410 HisLysLeuLeuGlyArgThrGlyIleGlyLysValPheGlyValValLysGlnIle 429
Db 1210 CACAAGCTCTTGGGACGACCTGGTGCATGATGAAGTTCTTGGTGTGGTGAATCAGATT 1269
Qy 430 LysProValIlePheThrValValGluGlnGlnSerAsnHisAsnGlyProValPheLeu 449
Db 1270 AAACCGGAGATTTTCACTGTGTGTGAGCAGGAATCGAACCAATAATAGTCCGATTTCTTA 1329
Qy 450 AspArgPheThrGluSerLeuHisTyrThrSerThrLeuPheAspSerLeuGluGlyAla 469
Db 1330 GATCGTTTACTGAGTGGTTCATATTTACTTCGACGTTGTTTGTGCTGTTGAAGGTGTA 1389
Qy 470 ProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeu 489
Db 1390 CCGAGTGTCAAGACAAGGTCATGTCGAGGTTTTTACTTTGGGTAAACACAGATCTGCAACGTT 1449
Qy 490 ValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAsn 509
```

```
Db 1450 GTGCTTGTGATGGACCTGACCGGTTGAGCGTCATGAAACGTTGAGTCAGTCGAGGAAC 1509
Qy 510 ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAla 529
Db 1510 CGGTTCGGGCTGTGCGGCTGCGAGTATGTTGCAATGCGTTTAAGCAAGCG 1569
Qy 530 SerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGly 549
Db 1570 AGTATGCTTTTGGCTCTGTTCAACGGCGGTGAGGGTTATCGGTTGAGGAGATGACGGC 1629
Qy 550 CysLeuMetLeu 553
Db 1630 TGTCTCATGTTG 1641

RESULT 7
US-09-117-853-9
; Sequence 9, Application US/09117853
; Patent No. 6307126
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; EARLIER APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-117-853-9

Alignment Scores:
Pred. No.: 1.27e-212 Length: 1642
Score: 1929.00 Matches: 400
Percent Similarity: 76.95% Conservative: 34
Best Local Similarity: 70.92% Mismatches: 52
Query Match: 65.39% Indels: 79
DB: 3 Gaps: 9

US-10-030-194A-4 (1-572) x US-09-117-853-9 (1-1642)
Qy 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 151 ATGAAGAGAGATCATCATCAT-----CATCAT----- 177
Qy 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
Db 178 -----CAAGATAAGAAGACTATGATGATGAATGAAGAA 210
Qy 41 AspAspGluLeuLeuGlyValLeuLeuGlyTyrLysValArgSerSerGluMetAlaGluVal 60
Db 211 GACGAC-----GGTAACGGCATCGATGTT 234
Qy 61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db 235 GCTCAGAAACTCGAGAGCTTGAAGTTATGATGATGTTCAAGAGAGCATCTTTCT 294
Qy 81 HisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTrpLeuAspAsn 100
Db 295 CAACTCGCTACTGAGACTGTTTCACTATAATCCGGCGGAGCTTTACACGCTGGCTTGAATTC 354
Qy 101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db 355 ATGCTCCACCGACCTTAATCTCCCTCGTCT-----AACCGCGAG 393
```

```
Qy 121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
Db 394 TAC-----GATCTTAAAGCTATTTCCCGGT 417
Qy 141 AsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerSer----- 158
Db 418 GACGCGATTCTC-----AATCAGTTCGCTATCGATTCCGGCTTCTTCGCTCTAACCAA 468
Qy 159 -----AsnLysArgLeuLysProSerSerSerSerSerProAsp 169
Db 469 GCGCGCGGAGAGATACGTATACAAACAAGCGGTTGAAATGCTCAAAAC----- 519
Qy 170 SerMetValThrSerProAlaGlyValIleGlyThrThrValThrValThr 189
Db 520 -----GGCGTCGTGGAAACCACC-----ACAGCGACGGCT 549
Qy 190 GluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHis 209
Db 550 GAGTCAACTCGGCA-TGTGTCTCTGGTTGACTCGCAGGAGAACGGTGTGCGTCTCGCTTCAC 608
Qy 210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229
Db 609 GCGCTTTTGGCTTGGCTGAAGCTGTTCAAGAGAGAACTCTGACTGTGCGGGAAGCTCTG 668
Qy 230 ValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr 249
Db 669 GTGAGCAATCGGATTCTTAGCTGTTTCTCAATCGAGCTATGAGAAAGTCGCTACT 728
Qy 250 TyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProProGlnThrGlnIle 269
Db 729 TACTTCGCCGAAGCTCTCGCGCGCGGATTTACCGTCTCTCTCGCTCGCAGAGTCCAATC 788
Qy 270 AspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLys 289
Db 789 GACCACCTCTCTCCGATACTCTTCAGATGCACCTTCTACGAGACTTGTCTTATCTCAAG 848
Qy 290 PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgVal 309
Db 849 TTCGCTCATTACGGCGCAATCAAGCGATTCTCGAAGCTTTTCAAGGGAGAAAGAGTT 908
Qy 310 HisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeu 329
Db 909 CATGTCATTGATTCTCTATGAGTCAAGGCTCTTCAATGGCGCGGCTTATGACGCTCTT 968
Qy 330 AlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAla 349
Db 969 GCGCTTCGACTGTGTGCTCTCTCTTTCGGTTAACCGAATTTGGTCCACCGGACCG 1028
Qy 350 AspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle 369
Db 1029 GATAATTTTCGATTATCTTCATGAAGTTGGGTGAAGCTGGCTCATTTAGCTGAGCGGATT 1088
Qy 370 HisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSer 389
Db 1089 CACGTTGAGTTTGAAGTACAGAGGATTTGTGCTAACACTTTAGCTGATCTTGATGCTTCG 1148
Qy 390 MetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLeu 409
Db 1149 ATGCTTGAGCTTTAGACCAAGTGAATGTAATCTGTCGGTTAACTCTGTTTTCAGCTT 1208
Qy 410 HisLysLeuLeuGlyArgThrGlyIleGluLysValPheGlyValValLysGlnIle 429
Db 1209 CACAAGCTCTTGGGACACCTGGTCCGATCGATAAGGTTCTTGTGTGGTGAATCAGATT 1268
Qy 430 LysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPheLeu 449
Db 1269 AAACCGGAGATTTTCACTGTGTTGAGCAAGGAATCGAACCACATAATAGTCGATTTCTTA 1328
Qy 450 AspArgPheThrGluSerLeuHisTyrSerThrLeuPheAspSerLeuGluGlyAla 469
Db 1329 GATCGGTTTACTGAGTCTGTCATATTATCTCGACGTTGTTTGTACTCGTGGAGAGGTGTA 1388
Qy 470 ProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeu 489
```

```
Db 1389 CCGAGTGGTCAAGACAGGTCAATGTCGGAGGTTTACTTGGGTAACAACATCGCAACGTT 1448
Qy 490 ValAlaCySGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAsn 509
Db 1449 GTGGCTGTGTGATGAGCACTGACCGAGTTGAGCGCTCATGAACGCTGAGTCAGTGGAGGAAC 1508
Qy 510 ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAla 529
Db 1509 CGGTTCCGGTCTCTGCTGGGTTTCCGGCTGCACATATTTGGTTCGAATCGCGTTTAAAGCAAGCG 1568
Qy 530 SerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnGly 549
Db 1569 AGTATGCTTTTGGCTCTGTTCACCGCGGTGAGGTTATCGGGTGAGGAGAGTGCAGGC 1628
Qy 550 CysLeuMetLeu 553
Db 1629 TGTCTCATGTTG 1640

RESULT 8
US-09-911-154-9
; Sequence 9, Application US/09911154
; Patent No. 6478809
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,154
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-154-9

Alignment Scores:
Pred. No.: 1,278-212 Length: 1642
Score: 1929.00 Matches: 400
Percent Similarity: 76.95% Conservative: 34
Best Local Similarity: 70.92% Mismatches: 52
Query Match: 65.39% Indels: 79
DB: 4 Gaps: 9

US-10-030-194A-4 (1-572) x US-09-911-154-9 (1-1642)
Qy 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 151 ATGAAGAGAGATCATCAT- 177
Qy 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
Db 178 -----CAAGATAAGAAGACTATGATGATGAATGAAGAA 210
Qy 41 AspAspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVal 60
Db 211 GACGAC-----GATAACCGCATGATGTT 234
Qy 61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db 235 GCTCAGAAACTCGACGACCTTGAAGTTATGATGCTTAATGTTCAAGAGAGATCTTCT 294
Qy 81 HisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTrpLeuAspAsn 100
```

```
Db 295 CAACTCGCTACTGAGACATGTTCACTATATTCGGCGGAGCTTTACACGTGGCTGATTCT 354
Qy 101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db 355 ATGCTACCGACCTTAATCTCCGTCGTC- 393
Qy 121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
Db 394 TAC-----GATCTTAAAGCATATCCCGGT 417
Qy 141 AsnAlaValCysArgArgSerAsnGlnPheAlaValAspSerSerSer----- 158
Db 418 GACGCGATTCTC-----AATCAGTTCGCTATCGATTTCGCTTCTTCGCTTAACCAA 468
Qy 159 -----AsnLysArgLeuLeuLysProSerSerSerProAsp 169
Db 469 GCGCGCGGAGGAGATACGTATACTACAACAACGCGTTGAAATGCTCAAC----- 519
Qy 170 SerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValThr 189
Db 520 -----GGCGTCGTGGAAACCAAC- 549
Qy 190 GluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHis 209
Db 550 GAGTCAACTCGGCA-TGTGTCCTCGTTGACTCGCAGGAGAACGCGTGCCTCTCGTTCCAC 608
Qy 210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229
Db 609 CGCCTTTTGGCTTGGCTGGAAGCTGTTCAAGAGGAGAAATCGACTGTCGCGGAGACTCTG 668
Qy 230 ValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr 249
Db 669 GTGAAGCAAAATCGGATCTTAGCTGTTTCTCAATCGGAGCTATGAGAAAAGTCTACT 728
Qy 250 TyrPheAlaGluAlaLeuAlaArgIleTyrArgLeuSerProGlnThrGlnIle 269
Db 729 TACTTCGCGGAAGCTCTCGCGCGGAGTTTACCCTCTCTCTCCGTCGACAGAGTCCCAATC 788
Qy 270 AspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLys 289
Db 789 GACCACCTCTCTCGATCTCTCAGATGACCTTCTACGAGACTTGTCTTATCTCAAG 848
Qy 290 PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysArgVal 309
Db 849 TTCGTCACCTTCACGGCGAATCAAGCGATTCCTGAAGCTTTTCAAGGGAAGAAAAGAGTT 908
Qy 310 HisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeu 329
Db 909 CATGTCATTGATTTCTCTATGAGTCAAGGCTTCAATGCGCGCGCTTATGACGCTCTT 968
Qy 330 AlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAla 349
Db 969 GCGTTCGACCTGTGTGTCCTCTCTGTTTCCGGTTAACCGGAATTTGGTCCACCGGACCG 1028
Qy 350 AspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle 369
Db 1029 GATAATTTGATTTATCTCATGAAGTTGGGTGAAGCTGGCTCATTTAGCTGAGGCGATT 1088
Qy 370 HisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSer 389
Db 1089 CACGTTGAGTTTGAGTACAGAGGATTTGTGGCTAACACTTTAGCTGATCTTGTGCTTCG 1148
Qy 390 MetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLeu 409
Db 1149 ATGCTTGAGCTTAGACCAAGTGAGATTGAATCTGTTCGGGTTAACTCTGTTCAGGCTT 1208
Qy 410 HisLysLeuLeuGlyArgThrGlyIleGluLysValPheGlyValValLysGlnIle 429
Db 1209 CACAAGCTCTTGGGACGACCTGGTCCGATCGATAAGGTTCTTGGTGTGGTGAATCAGATT 1268
Qy 430 LysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPheLeu 449
```



```
QY 350 AspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle 369
|||||
Db 1029 GATAATTCGATTATCTTCATCAAGCTGGTGAAGCTGGCTCATTTAGCTGAGCGGATT 1088
QY 370 HisValGluPheGluTyArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAsaSer 389
|||||
Db 1089 CACGTTGAGTTTGAAGTACAGAGGATTTGGGCTAACACTTTAGCTGATCTTGCCTTCG 1148
QY 390 MetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLeu 409
|||||
Db 1149 ATGCTTGAGCTTAGACCAAGTGAAGTGAATCTGTTGGGTTAACTCTGTTTCAGGCTT 1208
QY 410 HisLysLeuLeuGlyArgThrGlyGlyLeuLysValPheGlyValValLysGlnIle 429
|||||
Db 1209 CACAAGCTCTTGGGACGACCTGGTCCGATCGATAAGGTTCTTGGTGTGCGTAATCAGATT 1268
QY 430 LysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPheLeu 449
|||||
Db 1269 AAACCCGAGATTTTCACTGTGGTTGAGCAGGGAATCGAACCAATAATAGTCCGATTTCCTTA 1328
QY 450 AspArgPheThrGluSerLeuHisTyTyTrSerThrLeuPheAspSerLeuGluGlyAla 469
|||||
Db 1329 GATCGTTTACTGAGCTCGTTGCAATTATCTCGACGTTGTTGACTCGTTGGAAGGTGTA 1388
QY 470 ProSerSerGlnAspLysValMetSerGluValTyLeuGlyLysGlnIleCysAsnLeu 489
|||||
Db 1389 CCGAGTGTGTCAGACAAGCTCATGTCCGAGGTTTACTTGGGTAAACACATCTGCAACGTT 1448
QY 490 ValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAsn 509
|||||
Db 1449 GTGGCTTGTGAGTGGACCTGACCGAGTTGAGCGTCAATGAACGTTGAGTGGAGGGAAC 1508
QY 510 ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAla 529
|||||
Db 1509 CGTTTCGGGCTCGCTGGGTTTCCGGCTGCACATATGTTGCAATCGCTTTAAGCAAGCG 1568
QY 530 SerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyArgValGluLysAsnAsnGly 549
|||||
Db 1569 AGTATGCTTTTGGCTCTGTTCACCGCGTGAGGCTTATCGGGTGCAGGAGAGTGCACGCG 1628
QY 550 CysLeuMetLeu 553
|||||
Db 1629 TGTCTCATGTTG 1640
```

RESULT 12

```
US-09-911-514-5
; Sequence 5, Application US/09911514
; Patent No. 6794560
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jintong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911.514
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-514-5
```

Alignment Scores:

```
Pred. No.: 6,27e-212 Length: 1642
Score: 1923.00 Matches: 399
Percent Similarity: 76.95% Conservative: 35
Best Local Similarity: 70.74% Mismatches: 52
Query Match: 65.19% Indels: 79
DB: 4 Gaps: 9
US-10-030-194A-4 (1-572) x US-09-911-514-5 (1-1642)
QY 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 151 ATGAAGAGAGATCATCATCAT-----CATCAT----- 177
QY 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
Db 178 -----CAAGATAAGAAGACTATGATGATGAATGAAGAA 210
QY 41 AspAspGluLeuLeuGlyValLeuGlyTyTrLysValArgSerSerGluMetAlaGluVal 60
Db 211 GACGAC-----GGTAACGGCATGATGTT 234
QY 61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db 235 GCTCAGAAACTCGAGCAGCTTGAAGTTATGATGTTCAATGTTCAAGAGACGATCTTTCT 294
QY 81 HisLeuAlaThrAspThrValHisTyTrAsnProAlaGluLeuTyTrSerTrpLeuAspAsn 100
Db 295 CAACCTCGCTACTGAGACTGTTCACTATAATCCGCGGAGCTTTACACGCTGGCTTGATTCT 354
QY 101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db 355 ATGCTCACCGACCTTAATCTCCCTCGCTCT-----AACCCCGAG 393
QY 121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
Db 394 TAC-----GATCTTAAGACTATTTCCCGT 417
QY 141 AsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerSer----- 158
Db 418 GACGCGATTCTC-----AATCAGTTCTGCTATCGATTTCGGCTTCTTCGCTCAACCAA 468
QY 159 -----AsnLysArgLeuLysProSerSerSerSerProAsp 169
Db 469 GCGCGCGGAGGAGATACGTATACTACAAACAGCGGTTGAAATGCTCAAC----- 519
QY 170 SerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValThr 189
Db 520 -----GGCGTCTGGGAAACCAACC-----ACACGCGAGCGT 549
QY 190 GluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHis 209
Db 550 GAGTCAACTCGGCATGTTGTCCTGCTGACTCGCAGGAGAACGGTGTGCGTCTCGTTTCAC 609
QY 210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229
Db 610 CGCGTTTTGGCTTGGCTGGAAGCTGTTTCAGAGAGGAGAACTCGACTCGCGCGGAACCTG 669
QY 230 ValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr 249
Db 670 GTGAAGCAAAATCGGATTTCTTAGCTGTTTCTCAATCGGAGCTATGAGAAAGTCCCTACT 729
QY 250 TyrPheAlaGluAlaLeuAlaArgArgIleTyArgLeuSerProProGlnThrGlnIle 269
Db 730 TACTTCGCCGGAAGCTCTCGCGCGCGGATTTTACCGTCTCTCTCCGTCGACAGTCCAATC 789
QY 270 AspHisSerLeuSerAspThrLeuGlnMetHisPheTyTrGluThrCysProTyTrLysLys 289
Db 790 GACCACCTCTCTCCGATCTCTTTCAGATGCACTTCTACGAGACTTGTCTTATCTCAAG 849
QY 290 PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgVal 309
Db 850 TTCGCTCACTTTCACGCGGCAATCAACCGGATTTCTCGAAGCTTTTTCAGGGAAGAAAGATT 909
```

```
Qy 310 HisValIleAapPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeu 329
Db 910 CATGTCATTGA-TTCTCTATGAGTCAAGGTCTTCAATGCCGCGCTTATGACAGCTCTT 968
Qy 330 AlaLeuArgGluGlyGlyProSerPheArgLeuThrGlyIleGlyProProAlaAla 349
Db 969 GCGCTTCGACCTGGTGGTCCCTCTGTTTCCGGTTAAACCGGAATTTGGTCCACCGCACCG 1028
Qy 350 AspAsnSerAspHisLeuHisGluValGlyCysValLeuAlaGlnLeuAlaGluAlaIle 369
Db 1029 GATAATTTTCGATTATTCATGAAGATTTGGGTGAAGCTGGCTCATTTAGCTGAGCGGATT 1088
Qy 370 HisValGluPheGluTyArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSer 389
Db 1089 CACGTTGAGTTTGTAGTACAGAGGATTTGTGGCTAACACTTTAGCTGATCTTTGATGCTTCG 1148
Qy 390 MetLeuGluLeuArgProSerGluThrGluAlaValAlaValAlaAsnSerValPheGluLeu 409
Db 1149 ATGCTTTGAGCTTAGACCAAGTGAGATTGAATCTGTTGCGGTTAACTCTGTTTTCGAGCTT 1208
Qy 410 HisValLeuLeuGlyArgThrGlyGlyIleGluValPheGlyValValValysGlnIle 429
Db 1209 CACAAGCTCTTGGGACGACCTGGTCGATCGATAAGGTTCTTGGTGTGGTGAATCAGATT 1268
Qy 430 LysProValIlePheThrValValGluGlnGlnSerAsnHisAsnGlyProValPheLeu 449
Db 1269 AAACCGGAGATTTTCACCTGTGGTTGAGCAGGAATCGAACCATATAGTCCGATTTTCTTA 1328
Qy 450 AspArgPheThrGluSerLeuHisTyTrpSerThrLeuPheAspSerLeuGluGlyAla 469
Db 1329 GATCGGTTTACTGAGTCGTTGCAATATTACTACGAGCTTGTGTTGACTCGCTTGGAAAGGTGA 1388
Qy 470 ProSerSerGlnAspLysValMetSerGluValTyLeuGlyValysGlnIleCysAsnLeu 489
Db 1389 CCGAGTGTGTCAAGCAAGAGTCATGTCGGAGGTTTACTTGGGTAAACAGATCTGCAACGTT 1448
Qy 490 ValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAsn 509
Db 1449 GTGGCTGTGATGGACCTGACCGAGTTCAGCGCTCATGAAACGTTGAGTCAGTGGAGGAAC 1508
Qy 510 ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheIysGlnAla 529
Db 1509 CCGTTCCGGTCTCTGCTGGTGTTCGCGCTGCACATATTGTTCCGAATGCGTTTAAAGCAAGC 1568
Qy 530 SerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyArgValGluLysAsnAsnGly 549
Db 1569 AGTATGCTTTTGGCTCTCTGTTCAACGCGGTGAGGTTTATCGGTTGAGGAGTGCAGCGC 1628
Qy 550 CysLeuMetLeu 553
Db 1629 TGTCTCATGTTG 1640
```

```
RESULT 13
US-09-117-853-7
; Sequence 7, Application US/09117853
; Patent No. 6307126
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; CURRENT FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1636
```

```
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-117-853-7

Alignment Scores:
Pred. No.: 7,52e-210 Length: 1636
Score: 1905.00 Matches: 397
Percent Similarity: 76.60% Conservative: 35
Best Local Similarity: 70.39% Mismatches: 52
Query Match: 64.58% Indels: 81
DB: 3 Gaps: 10

US-10-030-194A-4 (1-572) x US-09-117-853-7 (1-1636)
Qy 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 151 ATGAAGAGAGATCATCATCAT-:::--CATCAT----- 177
Qy 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
Db 178 -----CAAGATAAGAAGACTATGATGATGAATGAAGAA 210
Qy 41 AspAspGluLeuLeuGlyValLeuGlyTyrlsValArgSerSerGluMetAlaGluVal 60
Db 211 GACGAC-----GGTAACGGCATGGATGTT 234
Qy 61 AlaLeuLysLeuGluGlnLeuThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db 235 GCTCAGAAACTCGAGCAGCTTGAAGTTATGATGCTAATGTTCAAGAAAGACGATCTTCT 294
Qy 81 HisLeuAlaThrAspThrValHisTyAsnProAlaGluLeuTyrlsSerTrpLeuAspAsn 100
Db 295 CAACTCGCTACTGAGACTGTTCACTATAATCCGCGGAGCTTTACACGCTGGCTTGAATCT 354
Qy 101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db 355 ATGCTCACCGACCTTAATCTCTCCGCTCT-----AACGCCAG 393
Qy 121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
Db 394 TAC-----GATCTTAAAGCTATTTCCTCCGCT 417
Qy 141 AsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerSer 158
Db 418 GACGCGATTCTC-----AATCAGTTCGCTATCGATTCGGCTTCTTCGCTCTAAACCA 468
Qy 159 -----AsnLysArgLeuLysProSerSerSerSerProAsp 169
Db 469 GCGCGCGGAGGAGATACGTATPACTACAAACAAGCGTTGAAATGCTCAAC----- 519
Qy 170 SerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValThr 189
Db 520 -----GGCGTCTGGNAACCAACC-----ACAGCGACGCT 549
Qy 190 GluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHis 209
Db 550 GAGTCAACTCGGCATGTTGCTCTGTTGACTCGCAGGAGAACGGTGTGCGTCTCGTTCAC 609
Qy 210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229
Db 610 CGCTTTTGGCTTGGCTGAAGCTGTTTCAGAAAGAGAAATCTGACTGTGGCGGAAGCTCTG 669
Qy 230 ValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr 249
Db 670 GTGAAGCAAAATCGGATTTCTAGCTGTTTCTCAATCGGAGCTATGAGAAAGTGCCTACT 729
Qy 250 TyrPheAlaGluAlaLeuAlaArgArgIleTyArgLeuSerProProGlnThrGlnIle 269
Db 730 TACTTCGCGGAAGCTCTCGCGCGGGATTTACCGCTCTCTCTCGCTCGCAGAGTCCAATC 789
Qy 270 AspHisSerLeuSerAspThrLeuGlnMetHisPheTyrlsGluThrCysProTyrlsLys 289
Db 790 GACCACCTCTCTCCGATCTTTCAGATGCACTTCTACGAGACTTGTCTTATCTCAG 849
```

```
Qy 290 PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysArgVal 309
Db 850 TTCGCTCACTTACCGCGCAATCAAGCGATTCTCGAAGCTTTTCAAGGGAAGAAAGAGTT 909
Qy 310 HisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeu 329
Db 910 CATGTCATTGATTCTCTATGAGTCAAGGTCTTGG-----GGCCTTATGAGGCTCTT 962
Qy 330 AlaLeuArgGluGlyGlyProSerPheArgLeuThrGlyIleGlyProProAlaAla 349
Db 963 GCGCTTCGACCTGGTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1022
Qy 350 AspIleAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle 369
Db 1023 GATAATTTTCGATTATCTTCAAGGTTGGGTGTAAGCTGGCTCATTTAGCTGAGCGGATT 1082
Qy 370 HisValGluPheGluTyArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSer 389
Db 1083 CACGTTGAGTTTGAGTACAGAGGATTTTGTGCTAACACTTTAGCTGATCTTGTGCTTCG 1142
Qy 390 MetLeuGluLeuArgProSerGluThrGluAlaValAlaValAlaAsnSerValPheGluLeu 409
Db 1143 ATGCTTGAGCTTAGACCAAGTGAGATTGAATCTGTTGGGTTAACTCTGTTTTCGAGCTT 1202
Qy 410 HisLysLeuLeuGlyArgThrGlyGlyIleGluLysValPheGlyValValLysGlnIle 429
Db 1203 CACAAGCTCTTGGGACGACCTGTCGATCGATAAGGTTCTTGGTGTGCTGAATCAGATT 1262
Qy 430 LysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPheLeu 449
Db 1263 AAACCGGAGATTTTACATGCTGTTGAGCAGGAATCGAAACCAATAATAGTCCGATTTCTTA 1322
Qy 450 AspArgPheThrGluSerLeuHisTyTySerThrLeuPheAspSerLeuGluGlyAla 469
Db 1323 GATCGTTTACTGAGTCTGTCATATTACTCGAGTTGTTGACTCGTTCGAAGGTGTA 1382
Qy 470 ProSerSerGlnAspLysValMetSerGluValTyLeuGlyLysGlnIleCysAsnLeu 489
Db 1383 CCGAGTGTGCAAGACAGTCAAGTCAATGTCGGAGGTTTACTTGGGTAACAGATCTGCACGTT 1442
Qy 490 ValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAsn 509
Db 1443 GTGCTTGTGTGATGACACCTGACCGGAGTGTGAGCGTCATGAACGTTGAGTCAGTGAGGAA 1502
Qy 510 ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAla 529
Db 1503 CCGTTCCGGTCTGCTGGGTTTGGGCTGCATATATTGTTGCAATGCGTTTAAAGCAGCG 1562
Qy 530 SerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyArgValGluLysAsnGly 549
Db 1563 AGTATGCTTTTGGCTCTGTTCAACGGCGGTGAGGTTATCGGTGGAGGAGTGACGGC 1622
Qy 550 CysLeuMetLeu 553
Db 1623 TGTCTCATGTTG 1634

RESULT 14
US-09-911-154-7
; Sequence 7, Application US/09911154
; Patent No. 6478809
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; FILE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; CURRENT APPLICATION NUMBER: US/09/911,154
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
```

```
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-154-7

Alignment Scores:
Pred. No.: 7.52e-210 Length: 1636
Score: 1905.00 Matches: 397
Percent Similarity: 76.60% Conservative: 35
Best Local Similarity: 70.39% Mismatches: 52
Query Match: 64.58% Indels: 81
DB: 4 Gaps: 10

US-10-030-194A-4 (1-572) x US-09-911-154-7 (1-1636)

Qy 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 151 ATGAAGAGAGATCATCAT-----CATCAT----- 177
Qy 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
Db 178 -----CAAGATAAGAAGACTATGATGATGAATGAAGAA 210
Qy 41 AspAspGluLeuLeuGlyValLeuGlyTyryLysValArgSerSerGluMetAlaGluVal 60
Db 211 GACGAC-----GGTAACGGCATGGATGTT 234
Qy 61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db 235 GCTCAGAAACTCGAGCAGCTTGAAGTTATGATGTCTAATGTTCAAGAGACGATCTTCT 294
Qy 81 HisLeuAlaThrAspThrValHisTyAsnProAlaGluLeuTySerTrpLeuAspAsn 100
Db 295 CAACCTCGCTACTGAGACTGTTCACTATAATCCGGCGGAGCTTACACGCTGGCTTGAATTCT 354
Qy 101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db 355 ATGCTCACCGACCTTAATCTCCCTCGCTCT-----AACGCCGAG 393
Qy 121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
Db 394 TAC-----GATCTTAAAGCTATTCCCGGT 417
Qy 141 AsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerSer----- 158
Db 418 GACGCGATTCTC-----AATCAGTTCTGCTATCGATTCTGGCTTCTTCGCTCTAACCAA 468
Qy 159 -----AsnLysArgLeuLysProSerSerSerSerProAsp 169
Db 469 GCGCGCGGAGGAGATACGTATATACAAACAAGCGTTGAAATGCTCAAC----- 519
Qy 170 SerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThr 189
Db 520 -----GGCGTCTGGNAACCAACC-----ACAGCGACGGCT 549
Qy 190 GluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHis 209
Db 550 GAGTCAACTCGGCATGTGCTCTGTTGACTCGCAGGAGAACCGGTGTCGCTCTCGTTTCTCAC 609
Qy 210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229
Db 610 GCGCTTTTGGCTTCGCTGAGCTGTTTTCAGAAAGAGAAATTCGACTGTGCGGAGAGCTCTG 669
Qy 230 ValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr 249
Db 670 GTGAAGCAATCGGATTCTTAGCTGTTTCTCAATCGGAGCTATGAGAAAAGTCTGCTACT 729
```

Qy 250 TyrPheAlaGluAlaLeuAlaAArgIleTyrArgLeuSerProGlnThrGlnIle 269
Db 730 TACTTCGCGAAGCTCTCGCGGGGGAATTTACCGTCTCTCTCGTCGAGAGTCCATC 789
Qy 270 AspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLys 289
Db 790 GACCACCTCTCTCCGATCTCTTCAGATGCACTTCTACGAGACTTGTCCCTATCTCAAG 849
Qy 290 PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysArgVal 309
Db 850 TTCCTCATTTCAGCGGGAATCAAGCGAATTCGAAGCTTTCAAGGGAAGAAAGATT 909
Qy 310 HisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeu 329
Db 910 CATGTCATTGATTTCTCTATGAGTCAGGCTCTGG-----GGCCTATGCAGGCTCTT 962
Qy 330 AlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAla 349
Db 963 GCGCTTCGACCTGTGTGCTCTCTCTGTTTCGGTTAACCGGAATTTGGTCCACCGGCACCG 1022
Qy 350 AspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle 369
Db 1023 GATAAATTCGATATATCTTCATGAAGTTGGGTGAAGCTGGCTCATTTAGCTGAGCGCAAT 1082
Qy 370 HisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAlaSer 389
Db 1083 CACGTTGAGTTTGATGACAGAGGATTTGTGGCTAACACTTTAGCTGATCTTGATGCTTCG 1142
Qy 390 MetLeuGluLeuArgProSerGluThrGluAlaValAlaValaAsnSerValPheGluLeu 409
Db 1143 ATGCTTGAGCTTAGACCAAGTGAGATGAATCTGTTGCGTTAACTCTGTTTTCGAGCTT 1202
Qy 410 HisLysLeuLeuGlyArgThrGlyGlyIleGluLysValPheGlyValValValLysGlnIle 429
Db 1203 CACAAGCTCTTGGGACGACCTGGTGGCGATCATGAAGTTCTTGGTGTGGTGAATCAGATT 1262
Qy 430 LysProValIlePheThrValValGluGlnGlnSerAsnHisAsnGlyProValPheLeu 449
Db 1263 AACCGGAGATTTTCACTGTGTTGAGCAGGAATCGAACCATATAGTCCGATTTTCTTA 1322
Qy 450 AspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAla 469
Db 1323 GATCGGTTTACTGAGTCGTCGATTAATTAATCTCGACGTTGTGTTGACTCGTTGGAAGGTGA 1382
Qy 470 ProSerGlnAspLysValMetSerLuuValTyrLeuGlyLysGlnIleCysAsnLeu 489
Db 1383 CCGAGTGGTCAAGACAAGGTCTATGTCGGAGGTTTACTTTGGGTAAACAGATCTGCAACGTT 1442
Qy 490 ValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAsn 509
Db 1443 GTGGCTTGTGATGACCTGACCGAGTTGAGCGTCAATGAAAGTTGAGTCAGTGGAGGAC 1502
Qy 510 ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAla 529
Db 1503 CGGTTCCGGTCTGCTGGGTTTGGCGTGCATATATGTTGGAATCGCTTTAAAGCAAGCG 1562
Qy 530 SerThrLeuLeuAlaLeuPheAsnGlyClyGluGlyTyrArgValGluLysAsnAsnGly 549
Db 1563 AGTATGCTTTTGGTCTGTTTCAACGGCGGTGAGGGTTATCGGGTGGAGGAGTGACGGC 1622
Qy 550 CysLeuMetLeu 553
Db 1623 TGTCTCATGTTG 1634

RESULT 15

US-09-911-514-7
; Sequence 7, Application US/09911514
; Patent No. 6794560
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E

; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911.514
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-911-514-7
Alignment Scores:
Pred. No.: 7.52e-210 Length: 1636
Score: 1905.00 Matches: 397
Percent Similarity: 76.60% Conservative: 35
Best Local Similarity: 70.39% Mismatches: 52
Query Match: 64.58% Indels: 81
DB: 4 Gaps: 10
US-10-030-194A-4 (1-572) x US-09-911-514-7 (1-1636)
Qy 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 151 ATGAAGAGAGATCATCATCAT-----CATCAT----- 177
Qy 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetMetValLysGluGluGlu 40
Db 178 -----CAAGATGAAGAGACTATGATGATGATGAAGAAGAA 210
Qy 41 AspAspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVal 60
Db 211 GACGAC-----GGTAACGGCATCGATGTT 234
Qy 61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db 235 GCTCAGAAACTCGAGACGCTTGAAGTTATGATGTCTAATGTTCAAGAAGACGATCTTTCT 294
Qy 81 HisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTyrLeuAspAsn 100
Db 295 CAACCTCGTACTGAGACTGTTCACTAATCCGCGGAGCTTTACACGTGCTGATTTCT 354
Qy 101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db 355 ATGCTACCGACCTTAATCTCTCGTCTGCT-----AACGCCGAG 393
Qy 121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
Db 394 TAC-----GATCTTAAAGCTATTTCGCGT 417
Qy 141 AsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerSer 158
Db 418 GACCGGATTTCT-----AATCAGTTCTGATTCGATTCGGCTTTCTTCGCTTAACCAA 468
Qy 159 -----AsnLysArgLeuLysProSerSerSerProAsp 169
Db 469 GCGCGCGGAGGAGATACGTATACAAACAACGGTTGAATGCTCAAAC----- 519
Qy 170 SerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValThr 189
Db 520 -----GGCGTCTGGAAACCAACC-----ACAGCGACGGCT 549
Qy 190 GluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHis 209
Db 550 GAGTCAACTCGGCATGTTGCTCTGTTGACTCTCGGAGGAGACGGTGTGCTGCTGCTTAC 609
Qy 210 AlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeu 229

Job time : 198 secs

```
Db 610 GCGCTTTTGGCTGCGTGAAGCTTTTCAGAAAGGAGAATCTGACTGTGGCGGAAGCTCTG 669
Qy VallysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThr 249
Db 670 GTGAAGCAAAATCGGATTTCTAGCTGTTCTCAATCGGAGCTATGAGAAAGTCCGCTACT 729
Qy 250 TyrPheAlaGluAlaLeuAlaArgIleTyrArgLeuSerProGlnThrGlnIle 269
Db 730 TACTTCGCGAAGCTCTCGCGGGCGATTTACCGTCTCTCTCCGTCGCAGATCCAATC 789
Qy 270 AspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLys 289
Db 790 GACCACCTCTCTCCGACTCTTCAGATGCATCTTACGAGACTTGTCCCTTATCTCAAG 849
Qy 290 PheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysArgVal 309
Db 850 TTCGCTCACTTCACGCGCAATCAAGGATTTCTCGAAGCTTTCAAGGGAGAAAGAGATT 909
Qy 310 HisValIleAspPheSerMetAsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLeu 329
Db 910 CATGTCATTGATTTCTCTATGAGTCAAGTCTTGG-----GCGCTTATGCAGGCTCTT 962
Qy 330 AlaLeuArgGluGlyGlyProSerPheArgLeuThrGlyIleGlyProProAlaAla 349
Db 963 GCGCTTCGACCTGGTGGTCTCTCTGTTTTCGCGTTAACCGGAATTGGTCCACCGCACCG 1022
Qy 350 AspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIle 369
Db 1023 GATAATTTTCGATTATCTCATGAGATTTGGGTGTAAGCTGGCTCATTTAGCTGAGCGGATT 1082
Qy 370 HisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSer 389
Db 1083 CACGTTGAGTTTGAGTCAGAGGATTTGTGGCTAACACTTTAGCTGATCTTGATGCTTCG 1142
Qy 390 MetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLeu 409
Db 1143 ATGCTTGAGCTTACGACCAAGTGAGATGAATCTGTTGCGGTTAACTCTGTTTCGAGCTT 1202
Qy 410 HisLysLeuLeuGlyValGlyThrGlyIleGluLysValPheGlyValLysGlnIle 429
Db 1203 CACAAGCTCTTGGGACGACCTGGTCCGATCGATAAGTTCTTGGTGTGTTGATCAGATT 1262
Qy 430 LysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPheLeu 449
Db 1263 AAACCGGAGATTTTCACTGTGGTTGAGCAGGAATCGAACCAATAATAGTCCGATTTCTTA 1322
Qy 450 AspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAla 469
Db 1323 GATCGGTTTACTGAGTCGTCGATTATTACTCGACGTTGTTTGTGACTCGTTGGAAGGTGTA 1382
Qy 470 ProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeu 489
Db 1383 CCGAGTGTCAAGACAAGTCAATGTCGAGGTTTACTTGGGTAAACAGATCTGCAACGTT 1442
Qy 490 ValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTyrSerAsn 509
Db 1443 GTGGCTTGTGATGGACCTGACCGAGTTGAGCGTCAATAAAGCGTTCAGTTCAGTGGAGGAAC 1502
Qy 510 ArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAla 529
Db 1503 CGGTTCCGGGTCCTGCTGGGTTTCGCGCTGCACATATTGGTTCGAATGCGTTTAAAGCAAGCG 1562
Qy 530 SerThrLeuAlaLeuPheAsnGlyGlyGlyTyrArgValGluLysAsnAsnGly 549
Db 1563 AGTATGCTTTTGGCTCTGTTTCAACGCGGTGAGGGTTATCGGGTCGAGAGAGTACCGGC 1622
Qy 550 CysLeuMetLeu 553
Db 1623 TGTCTCATGTTG 1634
```

Search completed: November 3, 2004, 11:25:21

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2004, 04:05:16 ; Search time 6464 Seconds
(without alignments)
3224.555 Million cell updates/sec

Title: US-10-030-194A-4
Perfect score: 2950
Sequence: 1 MKRDLHQGNHGTSIAGS.....LSWHTPLITTSAMKLSAVH 572

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10030194/runat_01112004.184752.28649/app.query.fasta_1.711
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10030194 @CGN 1.1.6425 @runat_01112004.184752.28649 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2220	75.3	2164	3	CNS0A04H Arabidops
2	2053	69.6	1842	3	CNS0A03T6 Arabidops
3	1886	63.9	1628	3	CNS0A03DM Arabidops
4	1591	53.9	2490	3	AY103865 Zea mays
5	1551	52.6	2101	3	CNS0A0E1H Arabidops
6	1419.5	48.1	1889	3	CNS0A0YH Arabidops
7	1242	42.1	840	8	BH708476 BOMIM09TR
8	1127	38.2	680	6	CB256202 89-B01274
9	1113	37.7	918	7	CK242923 EST726560

10	1105	37.5	877	7	CO083083	GR_Ea47P
11	1087.5	36.9	897	7	CK260776	EST706854
12	1042.5	35.3	870	7	CO071558	GR_Ea290
13	1035	35.1	845	7	CO124730	GR_Eb07N
14	1025	34.7	843	7	CK242924	EST726561
15	1006	34.1	948	9	CG961705	MBEBB32TR
16	998	33.8	943	9	CL291417	ZMBEBB063
17	990.5	33.6	962	7	CK294932	EST757646
18	971.5	32.9	718	9	CR492225	Medicago
19	970	32.9	817	7	CO104064	GR_Eb003
20	955.5	32.4	770	7	CO112484	GR_Eb004
21	954	32.3	852	7	CG308821	OG1E76TH
22	952.5	32.3	863	7	CO118420	GR_Eb020
23	950.5	32.2	1084	5	BX834999	EX834999
24	948	32.1	912	6	CA276066	SCCSD109
25	938.5	31.8	760	6	CA922517	EST640235
26	927	31.4	846	6	CB664102	OSJNEB09K
27	908.5	30.8	806	7	CO129906	GR_Eb31G
28	905.5	30.7	837	9	CG320924	OG1A546TH
29	901.5	30.6	834	6	CD438776	EL01N0517
30	895.5	30.4	831	7	CO127315	GR_Eb11N
31	894	30.3	726	1	AJ794151	AJ794151
32	892.5	30.3	653	7	CF808358	psHB034KJ
33	881	29.9	1854	3	CNS0ADXU	Arabidops
34	871	29.5	904	7	CK244071	EST727708
35	870	29.5	859	7	CO076317	GR_Ea37J
36	865.5	29.3	692	6	CA933156	MTUSCS.P4
37	858.5	29.1	743	2	AW774515	EST333666
38	853.5	28.9	739	8	BH552965	BOHUA01TR
39	845.5	28.7	809	6	CD438764	EL01N0517
40	843.5	28.6	626	2	BE204113	EST396789
41	837.5	28.4	788	9	CG344010	OGWXX24TV
42	832	28.2	796	6	CB682141	OSJNEF09B
43	828.5	28.1	677	6	CA927192	MTU6GR.P6
44	819.5	27.8	769	8	BH482148	BOHGS06TF
45	813.5	27.6	748	9	CC705631	OGUGF80TV

ALIGNMENTS

RESULT 1
CNS0A04H

LOCUS
DEFINITION

CNS0A04H 2164 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLT382E02 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).

ACCESSION

BX819848

VERSION

HTC; GSI:CDNA.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope

Submitted (18-NOV-2003)

BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen)

full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

source

Location/Qualifiers
1..2164
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLT382E02"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
complement(1..2164)
/gene="At2g01570"

gene

ORIGIN

Alignment Scores:

Pred. No.: 1 726-230 Length: 2164
Score: 2220.00 Matches: 477
Percent Similarity: 84.59% Conservative: 28
Best Local Similarity: 79.90% Mismatches: 53
Query Match: 75.25% Indels: 43
DB: 3 Gaps: 9

US-10-030-194a-4 (1-572) x CNS0A84H (1-2164)

Qy 1 MetLysArgAspLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerIleAla 18
Db 154 ATGAAGAGAGATCATCCAAATTCAGGTGTCACACCGGAGCTTCTTCTCA 213
Qy 19 GlySerSerThrSerProAlaValPheGlyLysAspLysMetMetValLysGlu 38
Db 214 TCATCATCAATCTCT-----AAAGATAAGATGATGTTGTAAGAAA 255
Qy 39 GluGlu-----AspAspGluLeuLeuGlyValLeuGlyTyrIysVal 52
Db 256 GAAGAAGACGGTGGAGGTAACATGGACGACGAGCTTCTCGTGTGTTTACAAAGTT 315
Qy 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGly 72
Db 316 AGGTCTCGAGATCGCGAGGTGCTTTGAGACTCGAACATAGACGATGATGAGT 375
Qy 73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92
Db 376 AATGTTCAAGAAGATGGTTTATCTCATCTCGCGACGGATACCTGTTTCAATATGATCCGTCG 435
Qy 93 GluLeuTyrSerTrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
Db 436 GAGCTTTATCTTGGCTTGATAATATGCTCTGAGCTTAATCCCTCCTCTTCCGCGC 495
Qy 113 GlySerAsnAlaLeuAsnProGluIleAsnAsnAsnAsnSerPhe---PheThr 131
Db 496 AGTTCTAACGGTTTAGATCCGGTCTTCTTCGCGCGAGATTGTTGGTTCGCGCTTCG 555
Qy 132 GlyGlyAspLeuLysAlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAla 151
Db 556 GATTATGACCTTAAAGTCATCCCGGAAACGGGATT-----TATCAGTTTCCG 603
Qy 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
Db 604 ---GCGATTGATTCCTCGCTTCGTCGAATAATCAGAACAGCGTTTGAGATCATGCTCG 660
Qy 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182
Db 661 AGTCTTGATTCATGGTTTACATCGACTTCGACCGGTACGAGATTGGTGGAGTCATAGGA 720
Qy 183 ThrThrValThrThrValThr-----GluSerThrArgProLeu 195
Db 195

Db 721 ACGACGGTGCAGCAACACCAACCAACGACGCGCGGGTGAAGTCAACTCGTCTGTGT 780
Qy 196 IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
Db 781 ATCTGTTGATCTCGCAGAG-AACGGTGTCTGTTTAGTCCACGCGCTTATGCGCTTGTCGA 839
Qy 216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235
Db 840 GAAGCAATCCAGCAGACAATTTGACTCTAGCGAGCT-CTTGTAAGCAAAATCGGATGC 898
Qy 236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrThrPheAlaGluAlaLeu 255
Db 899 TTAGCTGTCTCAAGCGGAGCTATGAGAGAAGTGGCTACTTACTTCGCCGAAGCTTTA 958
Qy 256 AlaArg-ArgIleTyrArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAs 275
Db 959 GCGCGGGGGATCTACCGTCTCTCTCCGCCGCAAGATCAGATCGATCATTTCTCTCCGA 1018
Qy 275 pThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAl 295
Db 1019 TACTCTTCAGATGCATTTTTCGAGAGACTTGTCTTATCTTAGGTTTCGCCCC-TTTACGGC 1077
Qy 295 aAsnGlnAlaIleLeuGluAlaPheGluGlyLysArgValHisValIleAspPheSe 315
Db 1078 GAACCAAGCGATTCCTCGAAGCTTTGAAGGTAAAGAGAGAGTACC-GTCAATTGATTTCTC 1136
Qy 315 rMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGly 335
Db 1137 GATGAACCAAGGTCTTCAATGGCTGCATTTATCGAAGCTCTTCGGCTTCGAGAAAGG 1196
Qy 335 yProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLe 355
Db 1197 TCCTCCAACTTTCCGGTTAAACCGGAATTGGTCCACCGCGCGGATAATTCTGATCATCT 1256
Qy 355 uHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTy 375
Db 1257 TCATGAAGTTGGTGTAAATTAGCTCAGCTTCGCGAGCGGATTCACGTAGAAATTCGAATA 1316
Qy 375 rArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgP 395
Db 1317 CCGTGGATTCGTTGCTTAACAGCTTAGCGATCTCGATGCTTCGATGCTTAGAGCTTAGACC 1376
Qy 395 oSerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyAr 415
Db 1377 GAGCGATACGGAAGCTGTTGGGTGAACCTCTGTTTTGAGCTACATAAGCTCTTAGGTTCG 1436
Qy 415 gThrGlyGlyIleGluLys--ValPheGlyValValLysGlnIleLysProValIlePhe 434
Db 1437 TCCCGGTGGGATAGAGAAACGATTTCTCGCGCTTGTGAACACAGATTAAACCCGGTATTTC 1496
Qy 435 ThrValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGlu 454
Db 1497 ACGTGGTTAGCAAGATCGAACATACGACCGGTTTTCTTAGACCGGTTTACTGAA 1556
Qy 455 SerLeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAsp 474
Db 1557 TCGTTACATTATTTCGACTCTGTTGATTGTTGGAAGGAGTTCGCAATAGTCAAGAC 1616
Qy 475 LysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGly 494
Db 1617 AAAGTCATGCTCGAAGTTTACTTAGGGAAACAGATTCTGTAATCTGGTGGCTTGTGAAGGT 1676
Qy 495 ProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAsnArgPheGlySerSer 514
Db 1677 CTGACAGAGTCGAGAGACACGAACCGTTGAGTCAATGGGGAACCGGTTTGGTTCGTCC 1736
Qy 515 GlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAla 534
Db 1737 GGTTTAGCGCGGCACATCTTGGGTCTAAACGGCTTAAACGAGCAGATGATGCTTTGTCT 1796
Qy 535 LeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGlyCysLeuMetLeuSer 554
Db 1797 GTGTTTAATAGTGCACCAAGGTTTTCGTGTGAGGAGAGTAAATGGATGTTTGTATGTTGGGT 1856

```
QY 555 TrpHisThrArgProLeuLeuThrThrSerAlaTrpLysLeuSer 569
|||||
Db 1857 TGGCACACTCGTCCACTCATTACCACCTCGGCTTGGAACCTCG 1901

RESULT 2
CNS0A3T6/c
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTPGH672D12 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
BX828248
VERSION
BX828248.1
GI:42461714
KEYWORDS
HTC; GSLT cDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1842)
Castelli, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1842)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
1..1842
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTPGH672D12"
/tissue_type="Hormone Treated Callus"
/plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.: 2,23e-212 Length: 1842
Score: 2053.00 Matches: 417
Percent Similarity: 81.41% Conservative: 34
Best Local Similarity: 75.27% Mismatches: 52
Query Match: 69.59% Indels: 51
DB: 3 Gaps: 7

US-10-030-194A-4 (1-572) x CNS0A3T6 (1-1842)

QY 32 LysMetMetValLysGluGluAsp-----AspGluLeuLeuGlyVal 47
|||||
Db 1639 AAGACTATGATGATTATGAGAGACGACGCGTACCGCATCGATGAGCTTCTACTGTT 1580
|||||
QY 48 LeuGlyTyLysValArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeu 67
|||||

1579 CTTTGTACTAGTTAGGTCAATCCGTAATGGCTGATGTTGCTCATAACTCGACGAGCTT 1520
QY 68 GluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAlaHis-LeuAlaThrAspThrVa 87
|||||
Db 1519 GAAGTTATGATGTTCTAATGTTCAAGAAGACGATCTTTCTTTTACTCGTACTGACAGCTG 1460
QY 87 LHisTyLysProAlaGluLeuLysSerTrpLeuAspAsnMetLeuThrGluLeuAsnPr 107
|||||
Db 1459 TCACCTATAATCCGGCGAGCTTTACACGCTGCTTGAATTTATGTTTCCACGACCTTAATCC 1400
QY 107 cProAlaAlaThrThrGlySerAsnAlaLeuAsnProGluLeuLeuAsnAsnAsnAs 127
|||||
Db 1399 TCCG-----TCGTCATAAGCC 1385
QY 127 nSerPhePheThrGlyGlyAspLeuLysAlaIleProGlyAsnAlaValCysArgArgSe 147
|||||
Db 1384 CGAGTTCTAT-----CTTAAGACTATTCGCGTGACGCGATTCTC----- 1345
QY 147 rAsnGlnPheAlaPheAlaValAspSerSerSer----- 158
|||||
Db 1344 -AATCAGTTCTGCTATCGATTTCGGCTTCTTCGTTAAACCAAGCGCGGAGGAGATACGTA 1286
QY 159 -----AsnLysArgLeuLysProSerSerSerProAspSerMetValThrSerProSe 176
|||||
Db 1285 TACTACAAACCAAGCGGTTGAATGCTCAAC----- 1255
QY 176 rProAlaGlyValIleGlyThrThrValThrValThrGluSerThrArgProLeuI 196
|||||
Db 1254 -----GGCGTCGTGGATACCACT--ACAGCGATGTTGAGTCACTCGGCATGTTGT 1205
QY 196 eLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAlaGl 216
|||||
Db 1204 CTTGTTTCACTCGCAGGAGAACGGTGTGCTCTCGTTTTCACGCGCTTTTTCGTTGCGCTGA 1145
QY 216 uAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPheLe 236
|||||
Db 1144 AGCTGTTTCAGAAAGAGAAATCTGACTGTAGCGGAAGCTCTGGTGAAGCAAAATCGGAATCTT 1085
QY 236 uAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrTyPheAlaGluAlaLeuAl 256
|||||
Db 1084 AGCGTTTCTCATATCGGCGGATGAGAAAGTCGCTACTTACTTTCGCGGAGCTCTCGC 1025
QY 256 aArgArgIleTyArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAspTh 276
|||||
Db 1024 CGCGCGGATTTACCGCTCTCTCTCGTCGAGAGTCCAATCGACCACTCTCTCTCGGATAC 965
QY 276 rLeuGlnMetHisPheTyArgLeuThrCysProTyLeuLysPheAlaHisPheThrAlaAs 296
|||||
Db 964 TCTTCAGATGCACCTCTACGAGACTTGTCTTATCTCAAGTTTCGCTCACTTCACGCGGAA 905
QY 296 nGlnAlaIleLeuGluAlaPheGluGlyLysArgValHisValIleAspPheSerMe 316
|||||
Db 904 TCAGCGGATTTTCGAGGCTTTTCAAGGGAAGAAAGAGTTTCATGTTCATTTCTCTAT 845
QY 316 tAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyPr 336
|||||
Db 844 GAGTCAAGGTTCTCAATGCGCGGCTTATGCGGCTCTTTCGCTTCGACCTGTGTGTC 785
QY 336 cProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeuHi 356
|||||
Db 784 TCTCTTTTTCGGTTAACCGGAATTTGGTCCACCGGACCGGATAATTTTCGATATCTTCA 725
QY 356 sGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyAr 376
|||||
Db 724 TGAAGTTGGGTGAAGCTGGCTCATTTAGCTGAGCGGATTCACGTTGAGTTGAGTACAG 665
QY 376 gGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgProSe 396
|||||
Db 664 AGGATTTGGCTAACACTTTAGCTGATCTTGTGCTTCGATCTTTCGATCTTTCGATCTTTCG 605
QY 396 rGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArgTh 416
|||||
Db 604 TGAGATTGAATCTGTTGCGGTTAACTCTGTTTTCAGGCTTCAAGCTCTTCAGGACGACC 545
```

```
QY 416 rGlyGlyIleGluIysValPheGlyValIysGlnIleYsProValIlePheThrVa 436
Db 544 TGGTCGATCGAATAGGTTCTGGTGTGGTGAATCAGATTAAACCGGAGATTTTCACATGT 485
QY 436 lValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSerIe 456
Db 484 GGTTCAGCAGGAAATCGAACCATATAGTCCGATTTCTTAGATCGGTTTACTGAGTCGTT 425
QY 456 uHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLysVa 476
Db 424 GCATATTACTCGACGTTGTTTACTCGTTGGAAGGTGTACCGAGTGTGTCAGACAAGGT 365
QY 476 lMetSerGluValTyrLeuGlyIysGlnIleCysAsnLeuValAlaCysGluGlyProAs 496
Db 364 CATGTCGAGGTTTACTTTGGGTAAACAGATTCGAACGTTGTGGCTTGTGATGACCTGA 305
QY 496 pArgValGluArgHisGluThrLeuSerGlnTyrSerAsnArgPheGlySerSerGlyPh 516
Db 304 CCGAGTTCAGCGTCATGAACGTTGAGTCAGTGGAGAACCGGTTCCGGTCTGCTGGGTT 245
QY 516 eAlaProAlaHisLeuGlySerAsnAlaPheIysGlnAlaSerThrLeuLeuAlaLeuPh 536
Db 244 TCGCGCTGCACATATTGTTGGAATCGCTTTAAGCAAGCAGATATGCTTTTGGCTCTGTT 185
QY 536 eAsnGlyGlyGluGlyTyrArgValGluIysAsnAsnGlyCysLeuMetLeuSerTyrHi 556
Db 184 CAACGGCGGTGAGGGTTATCGGTGGAGGAGTACCGGCTGTCTCATGTGGGTGGCA 125
QY 556 sThrArgProLeuIleThrThrSerAlaTyrIysLeuSer 569
Db 124 CACAGACCGCTCATAGCCACTTCGGCTTGGAACTCTCC 85

RESULT 3
CNS0A9DM
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSUTLS52807 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
BX819972
Accession
BX819972.1 GI:42468420
HTG; GSUT cDNA.
KEYWORDS
SOURCE
Arabidopsis thaliana
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1628)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1628)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
_length
```

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

```
FEATURES
source
1..1628
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSUTLS552E07"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
/complement(1..1628)
/gene="At2g01570"

ORIGIN
Alignment Scores: 3.08e-194 Length: 1628
Prid. No.: 1886.00 Matches: 392
Score: 86.58% Conservative: 21
Percent Similarity: 82.18% Mismatches: 38
Best Local Similarity: 63.93% Indels: 26
Query Match: 3 Gaps: 6
DB:

US-10-030-194A-4 (1-572) x CNS0A9DM (1-1628)
QY 114 SerAsnAlaLeuAsnProGluIleAsnAsnAsnAsnSerPhe---PheThrGly 132
Db 6 TCTAACGGTTAGATCCGGTCTCTTCCTTCGCGGAGATTTGTGGTTTCGGCTTCGAT 65
QY 133 GlyAspLeuIysAlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAlaPhe 152
Db 66 TATGACCTTAAAGTCATTCCGGAAACGGCATT-----TATCAGTTTCCG--- 110
QY 153 AlaValAspSerSerSer-----AsnIysArgLeuIysProSerSerSer 167
Db 111 GCGATTGATCTTCGTCTTCGTAATAATCAGAAACAGCGTTCGTGAGATCATGCTCGAGT 170
QY 168 ProAspSerMetValThrSerProSerPro-----AlaGlyValIleGlyThr 183
Db 171 CTTGATTTATGGTTTACATCGACTTCGACGGGTACGCAGATTGGTGAGTTCATAGGAACG 230
QY 184 ThrValThrThrValThr-----GluSerThrArgProLeuIle 196
Db 231 ACGGTGACGACACACACACACACACACACACACACACACACACACACACACACACACAC 290
QY 197 LeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAlaGlu 216
Db 291 CTGGTTGACTCGCAAGAGACGGGTGTCGTTTAGTCCACGGCTTATGGCTTGTGCAGAA 350
QY 217 AlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValIysGlnIleGlyPheLeu 236
Db 351 GCATCCAGCAGACAACATTTGACTCTAGCGGAAGCTCTTGTGAAGCAATCGGATGCTTA 410
QY 237 AlaValSerGlnAlaGlyAlaMetArgIysValAlaThrTyrPheAlaGluAlaLeuAla 256
Db 411 GCTGTGTTCTAAGCCGGAGCTATGAGAAAAGTGGCTACTTCTTCGCCGAAGCTTTACGC 470
QY 257 ArgArgIleTyrArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAspThr 276
Db 471 CGGCGGATCTACCGTCTCTCTCCCGCCGAGAAATCAGATCGATCATTTGTCTCTCCGATCT 530
QY 277 LeuGlnMetHisPheTyrGluThrCysProTyrLeuIysPheAlaHisPheThrAlaAsn 296
Db 531 CTTCCAGATGCATTTTACGAGACTTGTCTTATCTTAAATTCGCTCACTTCCACGGCGAAC 590
QY 297 GlnAlaIleLeuGluAlaPheGluGlyIysIysArgValHisValIleAspPheSerMet 316
Db 591 CAAGCGATTTCTGAAGCTTTTGAAGGTAAGAAGAGAGTACACGTCATTGATTTCTCGATG 650
QY 317 AsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGlyPro 336
Db 651 AACCAAGGTTCTCAATGGCTTGCACCTTATGCAAGCTTTGCGCTTCGCAAGAGAGGCTCT 710
QY 337 ProSerPheArgLeuThrGlyIleGlyProAlaAlaAspAsnSerAspHisLeuHis 356
```

```
Db 711 CCAACTTTCGGGTTAACCGGAATTGGTCACCGCGCGGATAAATTCGATCATCTTCAT 770
Qy 357 GluValGlyCysLeuAlaGlnLeuAlaGlu-AlaIleHisVal-GluPheGluTyrA 376
Db 771 GAAGTGTGGTATATTAGTCAGCTTCGGATGGGATTCCTCGTAGGATCGAATACC 830
Qy 376 rGlyPheValAlaAenSerLeuAlaAapLeuAapAlaSerMetLeuGluLeuArgProS 396
Db 831 GTGGATTGGTCTAACAGCTTAGCCGATCTCGATGCTTCGATGCTTAGCTTAGACGA 890
Qy 396 rGluThrCluAlaValAlaAenSerValPheGluLeuHisValLeuGluArgT 416
Db 891 GCGATACGGAAGCTGTTCGGTGAATCTGTTTTGAGCTACATAAGCTCTTAGTCTGTC 950
Qy 416 hrGlyGlyLeuGluValPheGlyValValValGln-IleLysProValIlePheThr 435
Db 951 CCGGTGGATAGAGAAAGTTCTCGCGGTGTGAACAGGATTAACCGGTGATTTCAAG 1010
Qy 436 ValValGluGlnGluSerAenHisAenGlyProValPheLeuAapArgPheThrGluSer 455
Db 1011 GTGGTTGAGCAAGAAATCGAACCATACGACCGGTTTTCTTAGACCGGTTTACTGAAATCG 1070
Qy 456 LeuHis-TyrTyrSerThrLeuPheAapSerLeuGluGlyAlaProSerSerGlnAapLy 475
Db 1071 TTACATATATTATTCGACTCTGTTGATTCGTTGGAAGGAGTTCGGAATAGTCAAGCAA 1130
Qy 475 sValMetSerGluValTyrLeuGlyLysGlnIleCysAenLeuValAlaCysGluGlyPr 495
Db 1131 AGTCATGCTGAAGTTTACTTAGGGAACAGATTGTGAATCTGGTGGCTTGTGAAGGTCC 1190
Qy 495 oAapArgValGluArgHisGluThrLeuSerGlnTrpSerAenArgPheGlySerSerGl 515
Db 1191 TGACAGAGTCGAGACACGAAACGTTGAGTCAATGGGGAACCGGTTTGGTTGCTCGCG 1250
Qy 515 yPheAlaProAlaHisLeuGlySerAenAlaPheLysGlnAlaSerThrLeuLeuAlaLe 535
Db 1251 TTTAGCGCGGCACATCTTGGGCTTAACGCGTTTAAGCAAGCGAGTATGCTTTTGTCTGT 1310
Qy 535 uPheAenGlyGlyGluGlyTyrArgValGluLysAenAenGlyCysLeuMetLeuSerTr 555
Db 1311 GTTTNATAGTGCCAAAGTTATCTGTGTGAGAGAGATTAATGATGTTTGTGTTGGTTG 1370
Qy 555 pHisThrArgProLeuIleThrThrSerAlaTrpLysLeuSer 569
Db 1371 GCACACTCGTCCACTCATTACCACCTCCGCTTGGAAACTCTCG 1413

RESULT 4
AY103865 2490 bp mRNA linear HTC 16-OCT-2002
LOCUS Zea mays P0148437 mRNA sequence.
DEFINITION Zea mays P0148437 mRNA sequence.
ACCESSION AY103865
VERSION AY103865.1 GT:21206943
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2490)
Hayney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2490)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
```

```
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
source
1..2490
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:638758"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration with the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Alignment Scores:
Pred. No.: 9,7e-162 Length: 2490
Score: 1591.00 Matches: 345
Percent Similarity: 65.79% Conservative: 82
Best Local Similarity: 53.16% Mismatches: 122
Query Match: 53.93% Indels: 100
DB: 3 Gaps: 15

US-10-030-194A-4 (1-572) x AY103865 (1-2490)
Qy 1 MetLysArgAapLeuHisGlnPheGlnGlyProAenHisGlyThrSerIleAlaGlySer 20
Db 198 ATGAAGCGCGAG---TACCAAGACGCGCGGAGTGGCGGCAC-----ATGGGCTCC 248
Qy 21 SerThrSerSerProAlaValPheGlyLysAapLysMetMetVal----- 36
Db 249 TCC-----AAGCAAGATGATGGCGCGCGCGGAGCA 284
Qy 37 -----LysGluGluGluAap---AapGluLeuLeuGlyValLeuGlyTyrLysValArg 53
Db 285 GGGGAACAGAGGAGGAGGAGCGTGGATGAGTCTGTGGCGCGCTCGGTTACAAGTGGT 344
Qy 54 SerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGly--- 72
Db 345 TCGTCGATATGGCGGAGCTCGCGCAGAAAGTGGAGCAGCTCGAGATGCCATGGGAGT 404
Qy 73 -----AsnAlaGlnGluAapGlyLeuAlaHisLeuAlaThr 84
Db 405 GCGCGGTGGCGCGCGCGCGCTACCGCTGATGACGGGTTCGTGTGCGACCTCGCCACG 464
Qy 85 AspThrValHisTyrAsnProAlaGluLeuTyrSerTrpLeuAapAsnMetLeuThrGlu 104
Db 465 GACACGCTGCATACATCCCTCCGACCTGCTCGTCTCGTCTGGTTCGAGAGCATGCTGTCGAG 524
Qy 105 LeuAenProProAlaAla----- 110
Db 525 CTCACGCGCGCTCAGCGCGCTCCCGACCGCGCGCGCGCCCAAGGCTCGCGTCCACA 584
Qy 111 -----ThrThrGlySerAsnAla-----Leu-AsnProGl 120
Db 585 TCGTCCACCGTCACAAAGTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644
Qy 120 uIleAenAsnAenAenAenAenSerPhePheThrGlyGlyAspLeuLysAlaIlePro--- 139
Db 645 CGTGGACTCGTCCAGCAGTACTAC-----GCTCTGAGACCGATCCCTC 689
Qy 140 -----GlyAenAlaValCysArg-ArgSerAenGlnPheAlaPheAlaValAapSers 157
Db 690 CGCGGTGGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 734
Qy 157 erSerAsnLysArgLeuLysProSerSerSerProAapSerMetValThrSerProSerP 177
Db 735 GGGAGCCCAAGCGATGAGGACTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 794
```

Qy	177	roAlaGlyValIleGlyThrThrValThrThrValThrThrGluSerThrArg-----	193
Db	795	CATCCATGGATGGCGTGCACACTAGGAGCTCCGTGTCGAAGTGCGCCGCCGCGCAGCG	854
Qy	194	-----ProLeuIleLeuValAspSerGlnAspA	203
Db	855	AAGCATCCGGCGGGCCAAACGGGCCCGCGTGCCAGTGGTGGTGGACACGACGAGG	914
Qy	203	snGlyValArqLeuValHisAlaLeuMetAlaCySAIAgluaAlaValaGlnSerSerAsnL	223
Db	915	CCGGATCCGGCTCGTCAACGGCTGTGGGCTGCGCGAGGCGCTGCACGAGAGA	974
Qy	223	euthrLeuAlaGluAlaLeuVallysGlnIleGlyPheLeuAlaValSerGlnAlaGlyA	243
Db	975	TCTCTGGCGGAGCGCTGTAAGCAGATCCCACGTCTGGCTTCGTCGAGGCGGTG	1034
Qy	243	laMetArqLySValAlaThrTyPheAlaGluAlaLeuAlaArqArqIleTyArqLeuS	263
Db	1035	CCATGCCAAGGTCCGCCCTACTTTCGGCAGGCGCTTGGCCGCGGTATCGTTCTCC	1094
Qy	263	erProPro-----GlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMethisP	281
Db	1095	GCCGCCAACGGACAGCTCCTCTCGACGGCGCTTCGCCGACCTCTTGACCGCGACT	1154
Qy	281	hetYrGluThrCysProTyriLeuIlysPheAlaHispheThrAlaAsnGlnAlaIleLeuG	301
Db	1155	TCTACGAGTCTGCCCCCTACCTGAAGTTCGCCCACTTCACCGCAACACGACCATTCTCG	1214
Qy	301	IuaIaPheGluGlyLySLeYargValHisValIleAspPheSerMetAsnGlnGlyLeuG	321
Db	1215	AGGCTTTCCGGCGTGC CGCGGTCCACGTGTCGACTTCGGCATCAAGCAGGGGATGC	1274
Qy	321	InTrpProAlaLeuMetGlnAlaLeuAlaLeuArqGluGlyGlyProProSerPheArgL	341
Db	1275	AGTGCCGGCTCTTCTCACGGCCCTCGCCCTCGCGCTGGCGCCCCCGTCTGTTCCGC	1334
Qy	341	euthrGlyIleGlyProProAlaAlaAspAenSerAspHisLeuHisGluValGlyCysL	361
Db	1335	TCACCGCGTGGCGCGCGCAGCCGACGACGACGACGCGCTTCGACACGTTGGCGTGA	1394
Qy	361	YSLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyArqGlyPheValAlaA	381
Db	1395	AAC TTGCCAGTTCGCGCACNACTATCCGCTGGACTTCAGTAGCCGTGGCTCTGTCGCG	1454
Qy	381	snSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArqProSerGluThr-----	398
Db	1455	CCACGCTCGCGACACTGAGCCGTTCTGCTGCAACCGAGGGCGATGACACGGATGACG	1514
Qy	399	-----GluAlavalAlavalAsnSerValPheGluLeuHisLySLeuLeuGlyArgThrG	417
Db	1515	AGCCGAGGTGATCCGCTGAAC TCCGTGTTTCGAGCTGCACCGGCTTCTTCGCGACGCCG	1574
Qy	417	lyGlyIleGluLySvalPheGlyValVallysGlnIleLySProValIlePheThrValV	437
Db	1575	GTGCCCTCGAAGGTCTCTGGCACGTGCGCGCGGTGGCGCCGAGGATCGTGACCGTGG	1634
Qy	437	alGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSerLeuH	457
Db	1635	TCGAGCAGGAGGCCAACACAAC TCCGGCACGTTCTTCGACCGCTTCACCGAGTCGTGTC	1694
Qy	457	istYrTySerThrLeuPheAspSerLeuGluGlyAlaProSerSerGln-----	473
Db	1695	ACTACTACTCACCATGTTGATTCTTCGAGGCGCGCGCGCGGTCTCGGCCCATGCCA	1754
Qy	474	-----AspLysValMetSerGluValTyriLeuG	483
Db	1755	CCGACGCTCCCGCGCGCGCGCGCACGACAGTCACTCGGAGGTGTACCTCG	1814
Qy	483	lyLySGlnIleCysAsnLeuValAlaCysGluGlyProAspArgValIgluarHisGluT	503
Db	1815	GCCGGCAGATCTCAAAGTGTGGTGCACAGGCGCGGCGCGCAGCGCAGCGACGACGAGA	1874

Qy	503	hrLeuSerGlnTnpSerAsnArgPheClySerSerGlyPheAlaProAlaHisLeuClyS	523
Db	1875	CGTTGGCCCAAGTGGCGCAGCGGCTTCGGCGGCTTCGGCGCCGTCACCTGGGCT	1934
Qy	523	erAsnAlaPheLyGlnAlaserThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyra	543
Db	1935	CCAAATGCTCAAGACGACGAGCAGCGTCTCGGCGCTTCGGCGCGGCGACGGGTACA	1994
Qy	543	rgValGlyLeuAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeuLeuThr	563
Db	1995	GRTTGGAGGAGAGGACGGGTGCTGACCCCTGGGGTGGCATACGCGCCCGCTCATCGCCA	2054
Qy	563	hrSerAlaTrpLysLeuSerAla	570
Db	2055	CCTCGCGTGGCGCGTCCGCC	2077
RESULT 5			
CNS0AE1H		2101 bp	linear
LOCUS		Arabidopsis thaliana Full-length cDNA Complete sequence from clone	HTC 06-FEB-2004
DEFINITION		GSLTLS492F11 of Adult vegetative tissue of strain col-0 of	
DEFINITION		Arabidopsis thaliana (thale cress).	
ACCESSION		EX815282	
VERSION		BA815282.1	GI:42474569
KEYWORDS		HTC; GSLT_CDNA.	
SOURCE		Arabidopsis thaliana (thale cress)	
ORGANISM		Arabidopsis thaliana	
REFERENCE		Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
REFERENCE		rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
AUTHORS		1 (bases 1 to 2101)	
REFERENCE		Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,	
AUTHORS		Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,	
REFERENCE		Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.	
TITLE		Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:	
REFERENCE		A Combined Approach to Evaluate and Improve Arabidopsis Genome	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 2101)	
AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :	
REFERENCE		BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr	
COMMENT		- Web : www.genoscope.cns.fr)	
REFERENCE		The sequences are based on single pass reads.	
COMMENT		Life Technologies (a division of Invitrogen) members carried out	
REFERENCE		full-length libraries construction : Temple G.	
COMMENT		Genoscope members carried out sequencing and annotation : Castelli	
REFERENCE		V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,	
COMMENT		Schachter V., Weissenbach J., Salanoubat M.	
REFERENCE		URGV INRA : Clepet C., Caboche M.	
COMMENT		Annotation is based on the June 2003 version of the Arabidopsis	
REFERENCE		genome released by MIPS (Munich Information center for Protein	
COMMENT		Sequences). 5 prime and 3 prime are assembled with Phrap.	
REFERENCE		http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full	
COMMENT		length	
REFERENCE		http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.	
FEATURES		Location/Qualifiers	
source		1. .2101	
FEATURES		/organism="Arabidopsis thaliana"	
source		/mol_type="mRNA"	
FEATURES		/strain="Col-0"	
source		/db_xref="taxon:3702"	
FEATURES		/clone="GSLTSL492F11"	
source		/tissue_type="Adult vegetative tissue"	
FEATURES		/plasmid_type="PCMVSPORT_6"	
source		1. .2101	
FEATURES		/gene="At1g14920"	
ORIGIN			
Alignment Scores:			
Pred. No.:		1.71e-157	2101
Score:		1551.00	353

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation
Unpublished
2 (bases 1 to 1889)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

source

1..1889
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTPGH592B06"
/tissue_type="Hormone Treated Callus"
/plasmid="pCMVSPORT 6"
complement(1..1889)
/gene="At5g17490"

gene

ORIGIN

Alignment Scores:

Pred. No.: 3,24e-143 Length: 1889
Score: 1419.50 Matches: 306
Percent Similarity: 69.30% Conservatives: 71
Best Local Similarity: 56.25% Mismatches: 112
Query Match: 48.12% Indels: 56
DB: 3 Gaps: 11

US-10-030-194A-4 (1-572) x CNS0A0YT (1-1889)

Qy 29 GlyIysAspLysMetMetMetValIysGluGluAspGluLeuLeuGlyValLeu 48
Db 215 GGAGACGATAACATG-----GACGAGTTTCTTGTCTTTTG 250
Qy 49 GlyTyrIysValArgSerSerGluMetAlaGluValAlaLeuIysLeuGluGlnLeuGlu 68
Db 251 GGTACAGGTTTCGATCTTCAGACATGGCAGATGTCACAGAGCTTGAACAGCTTGA 310
Qy 69 ThrMetMetGlyAsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHis 88
Db 311 ATGGTCTTGTCATGATATGCTCTTCT---AGTAATGCTTCAATGACACCGTTTCAT 367
Qy 89 TyrAnProAlaGluLeuTyrSerTrpLeuAspAsnMetLeuThrGluLeuAsnPro 108
Db 368 TACAATCTTCTGATCTCTCGGTTGGGCTCAGACATGCTCTCGGATCTTAAT----- 421
Qy 109 AlaAlaThrThrGlySerAsnAlaLeuAsnProGluIleAsnAsnAsnAsnAsnSer 128
Db 422 -----TACTACCGGATCTTGACCCGAAACCGGATT----- 451
Qy 129 PhePheThrGlyGlyAspLeuIysAlaIleProGlyAsnAlaValCysArgSerAsn 148
Db 452 -----TGGCATCTGAGACCAATCAGACGACGATGAGTGTTCAGATAGCAAT 499
Qy 149 GlnPheAlaPheAlaValAspSerSerSerAnIysArgLeuIysProSerSerPro 168
Db 500 -----AGTAACAGCAACAGAGGATTCGACTCGGTCTTGGTGT 538

Qy 169 AspSerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrVal 188
Db 539 GACTCA---GTGACCAGC----- 553
Qy 189 ThrGluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuVal 208
Db 554 ---GAGTCAACTCGTTCCTGGTGTCTTATC-----GAGGAGACAGAGTGTAGACTCGTT 604
Qy 209 HisAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAla 228
Db 605 CAGCGCTAGTGGCTCGCCGAGCGGTTCTGAGTGTGAGATCTGAGCCTCGCGATGCT 664
Qy 229 LeuValIysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAla 248
Db 665 CTCGTCAAGCGCTGGGATTACTCGCGGCTTCTCAAGCCGAGCATGCGGAAAGTCGT 724
Qy 249 ThrTyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProGlnThrGln 268
Db 725 ACCTACTTCGCGAAGCCTAGCTCGTAATTTACCGGATTCATCTTCGCGCCGCCGCC 784
Qy 269 IleAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeu 288
Db 785 ATTGATCTCTTCGAGAGGATTTTCAGATGAACCTTCTACGACTGTGTCCCTACTCTG 844
Qy 289 LysPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArg 308
Db 845 AAATTCGCTCATTTACGCGCAATCAGCGGATTTCTAGAAGCTGTTACGACGTCGCTGTC 904
Qy 309 ValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAla 328
Db 905 GTACACGTAATCGATCTAGGCTTAATCAAGGTATGCAATGGCCGCGCTTAATCAAGCC 964
Qy 329 LeuAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAla 348
Db 965 TTAGCTCTCGACCCGCTGGTCCACCGCTGTTTCTCTCCCGCGCTGGGAATCCGCTG 1024
Qy 349 AlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAla 368
Db 1025 -----AATCGAAGAGGATTTCAAGAGTTAGGTTTGAAGCTAGCTAGCTGGCTCAAGCC 1078
Qy 369 IleHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAla 388
Db 1079 ATCGCGCTCGAATTCAAATTCAAATGT-CTAACCGACGAGAGGTTATCCGATTTAGAACC 1137
Qy 389 SerMetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGlu 408
Db 1138 GATATGTTCCAGACCCCGA---ACCGAATCGAGACTCTAGTGGTTAATTCGGTTTCGAG 1194
Qy 409 LeuHisLysLeuLeuGlyArgThrGlyGlyIleGlyLysValPheGlyValValIysGln 428
Db 1195 CTTCAACCGGTTTATCCCAACCCGTTTCGATCGAAGAGCTGTTAGCGACGTTAAGCG 1254
Qy 429 IleLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPhe 448
Db 1255 GTTAAACCGGCTCTCGTAACAGTGTGTGAAACAAGACGACCAACACGCTGAGCTTTTC 1314
Qy 449 LeuAspArgPheThrGluSerLeuHisTyrTyrSerThrThrLeuPheAspSerLeuGluGly 468
Db 1315 TTAGCCGGTTTAAAGAACGCGCTTCACTATTCTCGAGCTTGTTCGACTTCCTCGAAGAT 1374
Qy 469 AlaPro-----SerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIle 486
Db 1375 GGTGTTGTGATACCGAGTCAAGACCGAGTCATGTCCGAGGTTTACTTAGGAGACAGATA 1434
Qy 487 CysAsnLeuValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGln 506
Db 1435 TTGAATTTGGTGGCCGCGAAGGAGCGATAGGATCGAGCGACACGACGACGCTGCTCAG 1494
Qy 507 TrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPhe 526
Db 1495 TGGCGAAAACCGTATGGGATCCCGCGGTTTGACCCGGTTAACCTCGGATCAGACCGTTT 1554
Qy 527 LysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGlyTyrArgValGluLys 546

```

Db      1555 AACCAAGCGAGTTGCTATTGGCGTTATCTGCGGTGGAGATGGATACAGATGGAGGAG 1614
Qy      547 AnAnGlyCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAlaTrp 566
Db      1615 AACGACGGAAGCTAATGCTTCGCTGGCAACGAAACCTCTAATCGCTGCATCGCGGTGG 1674
Qy      567 LysLeuSerAla 570
Db      1675 AAACGAGCGCG 1686

RESULT 7
BH708476      840 bp      DNA      linear      GSS 20-FEB-2002
DEFINITION    BOMIM09TR BO 2.3 KB Brassica oleracea genomic clone BOMIM09,
              genomic survey sequence.
ACCESSION     BH708476
VERSION       BH708476.1 GI:18794703
SOURCE        GSS.
ORGANISM      Brassica oleracea
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE     1 (bases 1 to 840)
AUTHORS       Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE         Whole genome shotgun sequencing of Brassica oleracea
JOURNAL       Unpublished (2001)
COMMENT       Other_GSSs: BOMIM09TF
              Contact: Chris Town
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA.
              Tel: 301-838-3523
              Fax: 301-838-0208
              Email: cdtown@tigr.org
              DNA is from a doubled haploid provided by Tom Osborn.
              Seq primer: TR
              Class: sheared ends.
FEATURES      Location/Qualifiers
              source          1..840
              /organism="Brassica oleracea"
              /mol_type="genomic DNA"
              /strain="TOL000DH3"
              /db_xref="taxon:3712"
              /clone="BOMIM09"
              /clone_lib="BO 2.3 KB"
              /note="Vector: pHOs1; Site 1: BstXI; 2-3 kb sheared
              genomic DNA inserted into pHOs1 using BstXI linkers"
ORIGIN
Alignment Scores:
Pred. No.:      2,05e-124      Length:      840
Score:          1242.00      Matches:      255
Percent Similarity: 93.01%      Conservative: 11
Best Local Similarity: 89.11%      Mismatches: 11
Query Match:     42.10%      Indels:      11
DBs:            8      Gaps:      2

US-10-030-194a-4 (1-572) x BH708476 (1-840)

Qy      154 ValAspSerSerSerAsnLysArgLeuLysProSerSerProAspSerMetValThr 173
Db      7 GTTAAACAGCCAGACGACCAAGCGTTTGAAGTCATGCTCGAGCCCTGATTGCTGTGTTACA 66
Qy      174 SerProSerProAlaGlyValIleGlyThrValThrValThr---GluSerThr 192
Db      67 -----GGACAACGCGTAACAACCAACCACTATTGAGTCAACT 102
Qy      193 ArgProLeuIleValAspSerGlnAspGlyValArgLeuValHisAlaLeuMet 212
Db      103 CGGTCCATCGCTCGTGTGACCTCGCAAGAGAACGGTGTGGCTAGTCACCGGCTCATG 162
Qy      213 AlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGln 232

```

```

Db      163 GCCTGCGCGGAAGCTATCCAGAACCAACGATTGTCTATAGCGGAAGCTCTCTGTCAAGCAG 222
Qy      223 IleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrTyrPheAla 252
Db      223 ATTGGATTCTTGGCGGTGTCTCAAGCGGAGCCATGAGAAAGTGCGCAGCTACTTCGCC 282
Qy      253 GluAlaLeuAlaArgArgIleTyrArgLeuSerProProGlnThrGlnIleAspHisSer 272
Db      283 GAAGCTCTCGCGCGCGGATCTACCGCTCTCTCGCGCGCAGACTCAGATCGATCACTCT 342
Qy      273 LeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLysPheAlaHis 292
Db      343 CTCTCGGACAGCTCCAGATGCACTTCTACGAGAGCTGCCCTTACCTCAAGTTCCTCTAC 402
Qy      293 PheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgValHisValIle 312
Db      403 TTCAGCGCCACCAAGCCATCTCGAGGCTTTCGAAGGGAAGAGAGATCCACGTCATC 462
Qy      313 AspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArg 332
Db      463 GACTTCTCATGAATCAAGGCTTCAATGGCGCGCTTATGCAAGCCCTTCGCTTGAGA 522
Qy      333 GluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspHisSer 352
Db      523 GAAGGTGCTCTCTCTGTTTCCGGTTAAACCGGGATTGGTCTCCGCGCGGATAACTCC 582
Qy      353 AspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGlu 372
Db      583 GATCATCTCCATGAGTTGGGTGAAGTTAGCTAGCTCGCGGAAGCCATTACGCTCGAG 642
Qy      373 PheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAlaSerMetLeuGlu 392
Db      643 TTTGAGTATCGTGGCTTCGTGCTTAATAGCTTGGCCGATCTCGATCTTCGATGCTGAG 702
Qy      393 LeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeu 412
Db      703 CTTAGACCGAGTGAGATCGAAGCTGTGGCGGTAACTCTGTTTCGAGCTCACAAAGCTC 762
Qy      413 LeuGlyArgThrGlyGlyLeuLysValPheGlyValValLysGlnLeuLysProVal 432
Db      763 CTTGGCCGTACCGGTGGGTAGAGNAAGTTCTCGCGCTGGTGAACAGATTAA-CCGGTG 820
Qy      433 IlePheThrValValGlu 438
Db      821 ATTTTCACCGTGGTTGAG 838

RESULT 8
CB256202      680 bp      mRNA      linear      EST 06-NOV-2003
LOCUS         89-E012741-027-007-A23-T7R MP1Z-ADIS-027 Arabidopsis thaliana cDNA
DEFINITION    clone MP1Zp772A237Q 5-PRIME, mRNA sequence.
ACCESSION     CB256202
VERSION       CB256202.1 GI:32880975
KEYWORDS      EST.
SOURCE        Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE     1 (bases 1 to 680)
AUTHORS       Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
              Mitchell-Olds,T. and Weisshaar,B.
TITLE         Large-scale identification and analysis of genome-wide
              single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
JOURNAL       Genome Res. 13 (6), 1250-1257 (2003)
MEDLINE       22683290
PUBMED        12799357
COMMENT       Contact: Weisshaar B
              ADIS DNA core facility at MP1Z
              Max-Planck-Institute for Plant Breeding Research
              Carl-von-Linne Weg 10, 50829 Koeln, Germany
              Fax: 00492215062851

```

Email: weissaha@piz-koeln.mpg.de
 Insert Length: 680 Std Error: 0.00
 Plate: 7 row: A column: 23
 Seq primer: 17R; CTAATAGACTCTACTATAGGA.
 Location/Qualifiers
 1. 680
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /ecotype="Bi-2"
 /db_xref="GABI:593924"
 /db_xref="taxon:3702"
 /clone="MP12p772A237Q"
 /tissue_type="whole plant"
 /dev_stage="adult plant, mixed stresses"
 /lab_host="E. coli TOP10"
 /clone_lib="MP12-ADIS-027"
 /note="Vector: pSP011; Site 1: SalI; Site 2: NotI; cDNA library from Arabidopsis thaliana, accession Eifel-2; ten week old total plants grown under long-day conditions in soil, whole adult plants were treated for 24 hours with different stresses, (1) at 4M-0 C in the dark, (2), at 37 Grad C in the dark, (3) lying in the lab after removing from soil, (4) in the greenhouse after wounding leaves with a forceps, (5) in the lab watering with a 150 mM NaCl solution, (6) at 26 M-0C in the light/UV; equal quantities of stressed plant material were pooled; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites SalI-NotI, primer sites and orientation:
 T7-SalI-CCACGGTCGG-Prime-cDNA-polyA-CC-NotI-SP6; Note: Sequencing granted in the context of the GABI Arabidopsis Verbund I: Genetic Diversity, 'Establishment of high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection' PI: Bernd Weisshaar Sequence submission managed by RZPD/GABI-Primary database: http://gabi.rzpd.de. This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."

ORIGIN

Alignment Scores:
 Pred. No.: 5,176-112 Length: 680
 Score: 1127.00 Matches: 218
 Percent Similarity: 97.79% Conservative: 3
 Best Local Similarity: 96.46% Mismatches: 4
 Query Match: 28.20% Indels: 1
 DB: 6 Gaps: 0

US-10-030-194A-4 (1-572) x CB256202 (1-680)

Qy 281 PheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeu 300
 Db 2 TTTTACGAGACTGTCTTAATTCGCTCACTTCACGGCGAACCAAGCGATTCTC 61
 Qy 301 GluAlaPheGluGlyLysArgValHisValIleAspPheSerMetAsnGlnGlyLeu 320
 Db 62 GAAGCTTTGAAGTAAGAGAGAGACTACACGTCATTGATCTTCGATGACCAAGGTCTT 121
 Qy 321 GlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArg 340
 Db 122 CAATGGCTGCACCTATGCAAGCTCTCGCTTCGAGAAGGAGGTCTCCCAACTTCGG 181
 Qy 341 LeuThrGlyIleGlyProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCys 360
 Db 182 TTAACCGGAATTGGTCCACGGCGCGGATTAATCTGATCATCTTCATGAAGTGTGGTGT 241
 Qy 361 LysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAla 380
 Db 242 AAATTAGCTCAGCTTCGGGAGCGATTACAGTAGAATTCGAATACCGTGGATTGCTGT 301
 Qy 381 AsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAla 400
 Db 302 AACAGCTTAGCCGATCTCGATGCTTCGATGCTTAGCTTAGACCGAGCGATACGGAAGCT 361

Qy 401 ValAlaValAenSerValPheGluLeuHisLysLeuLeuGlyArgThrGlyGlyLeuGlu 420
 Db 362 GTTCGGGTGAACCTCTGTTTTGAGCTACATAAGCTCTTAGTCTGCCGGTGGATAGAG 421
 Qy 421 LysValPheGlyValValLysGlnIleLysProValPheThrValValGluGlnGlu 440
 Db 422 AAAGTTCTCGCGCTGTGTGAACACAGATTAAACCGGTGATTTTTCACGGTGTGAGCAAGAA 481
 Qy 441 SerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSer 460
 Db 482 TCGAACCATTAACGACCGGTTTTCTTAGACCGGTTTACTGAATCGTTACATTATTATTCG 541
 Qy 461 ThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLysValMetSerGluVal 480
 Db 542 ACTCTGTTGATTCGTTGGAAGGAGTTCGGAATAGTCAAGACAAAGTCATGCTGAAGATT 601
 Qy 481 TyrLeuGlyLysGlnIleCysAsnLeuVal-AlaCysGluGlyProAspArgValGluAr 500
 Db 602 TACTTAGGGAACAGATTGTGAATCTGTTGGGCTTGTGAAGTCTCTGACAGAGTCGAGAG 661
 Qy 500 gHisGluThrLeuSer 505
 Db 662 ACACGAACAGTTGAGT 677

RESULT 9
 CK242923 918 bp mRNA linear EST 30-JUL-2004
 LOCUS EST728560 potato callus cDNA library, normalized and full-length
 DEFINITION Solanum tuberosum cDNA clone POCA175 5' end, mRNA sequence.
 ACCESSION CK242923
 VERSION CK242923.1 GI:39787029
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)

ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 918)
 AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
 TITLE Generation of ESTs from potato callus tissue
 JOURNAL Unpublished (2003)
 COMMENT Other_ESTs: EST726561
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics Institute via http://genome.arizona.edu/orders/ .
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 source
 1..918
 Location/Qualifiers
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POCA175"
 /tissue_type="callus"
 /lab_host="DHL0B-Toma"
 /clone_lib="potato callus cDNA library, normalized and full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Solanum tuberosum var. Kennebec callus tissue grown on solid media."

ORIGIN
 Alignment Scores:
 Pred. No.: 2,846-110 Length: 918
 Score: 1113.00 Matches: 216
 Percent Similarity: 80.53% Conservative: 28
 Best Local Similarity: 71.29% Mismatches: 42
 Query Match: 37.73% Indels: 17
 DB: 7 Gaps: 1

```
US-10-030-194A-4 (1-572) x CK242923 (1-918)

QY 265 ProGlnThrGlnLeuSerAspThrLeuGlnMetHisPheTyrGluThr 284
DB 11 CCACAAGATTCATGATCATATACAGATGTTTACAAATGCATTTACAGAACT 70
QY 285 CysProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaLeuGluAlaPheGlu 304
DB 71 TGCCTTATCTCAAAATTCGCGATTTACTGCAAAATCAAGCGATTTCTTGAAGCGTTTACA 130
QY 305 GlyLysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAla 324
DB 131 GGTGTGAACAAGTTCATGCTGATTCATGCTGATTCAGCTTAAACACAGGGTATGCAATGGCGTCA 190
QY 325 LeuMetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIle 344
DB 191 CTTATGCAAGCTTTAACTTTAGCCCGGTGACCGCTGCAATTTAGACTCACCGGAATA 250
QY 345 GlyProProAla-AlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlu 364
DB 251 GGCCTCCACAACCGAGATAATACAGATGCGTTGCAACAAGTTGGATGGATGGCTCA 310
QY 364 nLeuAlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAl 384
DB 311 GTTAGCTGAACAATCGGAGTTGAATTCGAATTCAGGGGATTTGTTCCTAATTCATTAGC 370
QY 384 aAspLeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAlaValAlaValAs 404
DB 371 AGATCTTGATGCACTATACCTGATATAGGCAAGTGAATGAGCGAGTAGCTATATAA 430
QY 404 nSerValPheGluLeuHisLysLeuGluArgThrGlyGlyIleGluLysValPheGlu 424
DB 431 CTCGTGTTTGGAGCTTCATCGATGCTATCCCGCGCGGAGCAATCGAAAAAGTGTAAA 490
QY 424 yValValLysGlnIleLysProValIlePheThrValGluGlnGluSerAsnHisAs 444
DB 491 TTTCTATTAACAAGATTAACCCCGAGATCGTTACTCTGTTGAACAAGAGCGAATCATAA 550
QY 444 nGlyProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAs 464
DB 551 CCAGGGGTTTTCATGATGATGATTTAACGAAGCTTTGCAATTTACTCAACCATGTTTGA 610
QY 464 pSerLeuGluGlyAlaProSerSer----- 472
DB 611 TTCGTTAGAAAGCTCTGGGTCTTCATCTTCGGCTTCACCAACTGGGATTTCTTCCTCAACC 670
QY 473 -----GlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAs 488
DB 671 TCCTGTGAACAATCAGGATTTGGTGATGTGCGAGGTTTATTTAGGAGACAGATTTGTAA 730
QY 488 nLeuValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSe 508
DB 731 CGTGTGGCTGTGAAGGTTTCAGATCGAGTTGAACGGGATGAAACACACTGAATCAATGAG 790
QY 508 rAsnArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlu 528
DB 791 GGTTAGGATGAATCATCTGGGTTCGATCCGCTTCATCTGGGTTCAAATGGCTTCAAAACA 850
QY 528 nAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnAs 548
DB 851 AGCTTCCATGCTTTTAGCTCTGTCGCCGCGCGGATGGTTACAGAGTGGAGGAAACGA 910
QY 548 nGlyCys 550
DB 911 TGGGTGT 917

RESULT 10
LOCUS C0083083 877 bp mRNA linear EST 15-JUN-2004
DEFINITION GR_Ea47P23.r GR_Ea Gossypium raimondii cDNA clone GR_Ea47P23 3',
mRNA sequence.
ACCESSION C0083083
```

```
VERSION C0083083.1 GI:48752564
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 877)
AUTHORS Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 47 row: P column: 23.
FEATURES
source
1..877
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea47P23"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."
Alignment Scores:
Pred. No.: 1..97e-109 Length: 877
Score: 1105.00 Matches: 218
Percent Similarity: 82.94% Conservative: 25
Best Local Similarity: 74.40% Mismatches: 46
Query Match: 37.46% Indels: 4
DB: 7 Gaps: 2
US-10-030-194A-4 (1-572) x C0083083 (1-877)
QY 216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235
DB 2 GAAGCTGTTCAACAAGATAATCTTAAACTAGCTGATGATGATGATGATGATGATGATGATGAT 61
QY 236 LeuAlaValSerGlnAlaGlyValaMetArgLysValAlaThrTyrPheAlaGluAlaLeu 255
DB 62 CTTCCTTCATCAACAACCTGGTGTATGAGAAAGTTGCTACTTATTTTCTGCTGAACTTTA 121
QY 256 AlaArgGlyLeuTyrArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAsp 275
DB 122 GCTGGAAGAATTTATAGAATTTTCCACCACAGATTCA---CTTGATCCATCATATAATGAT 178
QY 276 ThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295
DB 179 AAGTTACAAAATTCCTCTATGAAACTTGTCTTATTTTGAATAATTTGCTCATTTTACAGCC 238
QY 296 AsnGlnAlaIleLeuGluAlaPheGluGlyLysArgValHisValIleAspPheSer 315
DB 239 AATCAAGCCATATTGGAAGCTTTTCAATGGCTAGTAGATTCATGTTATGATTTTGGG 298
QY 316 MetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGly 335
DB 299 CTAAAAACAAGGTATGCAATGCCAGCTTTAATGCAAGCACCTTGCAATTAAGACACAGCTGA 358
QY 336 ProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeu 355
```

```
Db 359 CCACCGGGGTTTCGATTGACCGGAATTGACCGCTCAACCGGATAAATCTGATGCGTTG 418
Qy 356 HisGluValGlyCysLeuLeuAlaGlnLeuAlaGluAlaLeuHisValGluPheGluTyr 375
Db 419 CAACAAAGTGGGTGGAAAGCTAGCTCAATTTGGCCGAAACGATCGGATCGAATTCGAGTTT 478
Qy 376 ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395
Db 479 CGGGGATTCGTGGCTAATAGTTAGCCGATCTCGAACCCGAAATGCTCGATATTCGTCT 538
Qy 396 SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArg 415
Db 539 CCCAGATGGAAGTAGTAGCGGTGAACGCTGTTTCGAGCTTCATCCCTTGTAGTCGA 598
Qy 416 ThrGlyGlyLeuGluLysValPheGlyValValLysGlnLeuLysProValIlePheThr 435
Db 599 CCGGTGGGATCGAAAAAGTTGTTTCCTCTATTAAAGCGATGAACCCCAAGATTGTCACG 658
Qy 436 ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455
Db 659 GTTGTGACAAAGAGCGAATCACACCGTCCGGTTTCTTAGACCGTTTACTGAAGCT 718
Qy 456 LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGly-----AlaProSerSer 472
Db 719 CTCATTATTATTCTACCTTTTCGACTCGTTGGAAGGTTCAGGGGTGGCGCCACCGAGT 778
Qy 473 GlnAspLysValMetSerGluValTyrLeuGlyLysGlnLeuCysAsnLeuValAlaCys 492
Db 779 CAAGACCTGGCTATCTCGAGTTACTTAGAAGACAGATTGTGAACGTTGTTGCTGTTGT 838
Qy 493 GluGlyProAspArgValGluArgHisGluThrLeuSer 505
Db 839 GAAGGATGGACCGAGTTGAACGACACGACCGCAATTGACT 877

RESULT 11
CK260776 897 bp mRNA linear EST 03-AUG-2004
LOCUS EST706854 potato abiotic stress cDNA library Solanum tuberosum cDNA
DEFINITION clone POAB559 5' end, mRNA sequence.
ACCESSION CK260776
VERSION 1
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
Buell,C.R., Hart,A., Ziemann,V., Karamycheva,S.A. and Baker,B.
1 (bases 1 to 897)
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other_ESTs: EST706855
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: ART TAG GTG ACA CTA TAG.
Location/Qualifiers
1. 897
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAB559"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/notice="Vector: pCMVSPORT6.1; Site:1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
```

four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:
Pred. No.: 1.66e-107 Length: 897
Score: 1087.50 Matches: 210
Percent Similarity: 85.02% Conservative: 34
Best Local Similarity: 73.17% Mismatches: 42
Query Match: 36.86% Indels: 1
DB: 7 Gaps: 1

US-10-030-194A-4 (1-572) x CK260776 (1-897)

```
Qy 183 ThrThrValThrValThrGluSerThrArgProLeuLeuLeuValAspSerGlnAsp 202
Db 3 ACCTCTATGTTAACAGATTTCATCAGCAGCTAGACTGTTGTTACTAGTTGATTCACAGAA 62
Qy 203 AsnGlyValArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsn 222
Db 63 ACTGGTGTTCGTCCTGTTCTATCATCTTAATGGCTGTGTAAGCTGTACAAACAAGAAAT 122
Qy 223 LeuThrLeuAlaGluAlaLeuValLysGlnLeuGlyPheLeuAlaValSerGlnAlaGly 242
Db 123 CTAACTCTAGCTGATCAACTTGTGTACATATGTTGTTGTTGTTTTCATCATCTGTT 182
Qy 243 AlaMetArgLysValAlaThrTyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeu 262
Db 183 GCTATGAGAAAGTTGTTACTTACTTCTGCTGAGCATTTAGCAAGAGATCTACAAAT 242
Qy 263 SerProProGlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHisPheTyr 282
Db 243 TAT---CCACAAGATTCAATTGAATCATCATACAGATGTTTACAAATGCATTTCTAC 299
Qy 283 GluThrCysProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAla 302
Db 300 GAAATCTGCCCTTATCTCAAATTCGCGATTTTACTGCAATCAAGCGATTTCTTGAAGCG 359
Qy 303 PheGluGlyLysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTyr 322
Db 360 TTTACAGTTGTACAAAGTTTCATGTCATGATTTTCAGTTTAAACACAGGGTATGCAATGG 419
Qy 323 ProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArgLeuThr 342
Db 420 CTTGCACCTTATGCAAGCTTTAGCTTTACGCCCGGTGGACCGCTGCAATTTAGACTCAC 479
Qy 343 GlyIleGlyProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeu 362
Db 480 GGAATAGCCCTCCACACCGGATAATACAGATCGTTGCAACAGTTGGATGGAGTTG 539
Qy 363 AlaGlnLeuAlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAsnSer 382
Db 540 GCTCAGTTAGCTGAACCAATCGGAGTTCAATTCGAATTCAGGGGATTTGTTGCTAATTC 599
Qy 383 LeuAlaAspLeuAspAlaSerMetLeuGluLeuTyrArgProSerGluThrGluAlaValAla 402
```

```

Db      600 TTACGAGATCTTGATGCAACTATATTGATATAGGCCAAGTGAACCTGAAGCAGTAGCT 659
Qy      403 ValAenSerValPheGluLeuHisLysLeuGlyArgThrGlyGlyLeuLysVal 422
      ::::|||||
Db      660 ATAAACTCTGTTTGGAGCTTCATGATGCTATCCCGCGGAGCAATCGAAAAGTG 719
Qy      423 PheGlyValValLysGlnIleLysProValIlePheThrValValGluGlnSerAen 442
      ::::|||||
Db      720 TTAATTCATTATAACAGATTAAACCCGAAGATCGTTACTTGTGTGAACAGAGCGAAT 779
Qy      443 HisAenGlyProValPheLeuAspArgPheThrGluSerLeuHisLysLysValSerThrLeu 462
      ::::|||||
Db      780 CATACGCGAGGGTTTTCATGTATGATAGATTAAAGAGCTTGTGCATATTACTCAACCATG 839
Qy      463 PheAspSerLeuGluGlyAla 469
Db      840 TTTGATTCGTTAGAAAGCTCT 860

RESULT 12
CO071558 870 bp mRNA linear EST 15-JUN-2004
LOCUS GR_Ea29004.r GR_Ea Gossypium raimondii cDNA clone GR_Ea29004 3',
DEFINITION mRNA sequence.
ACCESSION CO071558.1 GI:48741039
VERSION CO071558
KEYWORDS Gossypium raimondii
SOURCE EST.
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 870)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 29 row: O column: 04.

FEATURES
source
1..870
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea29004"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_libs="GR_Ea"
/notes="Vector: pCMV.SPORT-6.1; Site1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN
Alignment Scores:
Pred. No.: 1.29e-102 Length: 870
Score: 1042.50 Matches: 213
Percent Similarity: 79.52% Conservative: 20
Best Local Similarity: 72.70% Mismatches: 43
Query Match: 35.34% Indels: 17
DB: 7 Gaps: 4

US-10-030-194A-4 (1-572) x CO071558 (1-870)

Qy      136 LysAlaIleProGlyAsnAlaVal-----CysArg 145
      ::::|||||

```

```

Db      1 CAAGCCATTCTCGTAAGGCTATTATTATCTTAAATACCCAAACCCCAACCCAC 60
Qy      146 ArgSerAenGlnPheAlaValAspSerSerSerAenLysArgLeuLysProSer 165
      ::::|||||
Db      61 GATTCTCTCTCTTCCACACCCCAACCGTAAATCCGAAAACCGCTTCAATCTACA 120
Qy      166 Ser-----SerProAspSerMetValThrSerProSerProAlaGlyValIleGly 182
      ::::|||||
Db      121 TCCGGGCGCGCCTCGCCCTCGGATATCTTCCCCCTCTCTCTCCGCT-----GCT 171
Qy      183 ThrThrValThrValThrGluSerThrArgProLeuIleLeuValAspSerGlnAsp 202
      ::::|||||
Db      172 GCTTCTCTATGGAATTCGACATGATCAACTCGTCCGCTGCTCTGTTGACTCGCAAGAA 231
Qy      203 AsnGlyValArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAen 222
      ::::|||||
Db      232 AATGGAATCCGGCTTGTCCATGCTTTGATGCGATGTGCGGAAGCTGTCCAGCAAAACAAT 291
Qy      223 LeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPheLeuAlaValSerGlnAlaGly 242
      ::::|||||
Db      292 CTTAATCTGGCTGAAGCTTTAGTTAAACAAATGGGTTCTTAGCGATTTCTCAGGCTGGG 351
Qy      243 AlaMetArgLysValAlaThrTyPheAlaGluAlaLeuAlaArgArgIleTyArgLeu 262
      ::::|||||
Db      352 GCTATGAGAAAAGTGGCTACTTATTTCCGAGAGCATTTAGCCCGAAGATTTACAGATT 411
Qy      263 SerProGlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHisPheTy 282
      ::::|||||
Db      412 TAT--CCTAAAACCCCACTCGACCTCTTTTTCAGATGTTCTTCACATGCACCTTTTAC 468
Qy      283 GluThrCysProTyLeuLysPheAlaHisPheThrAlaAenGlnAlaIleLeuGluAla 302
      ::::|||||
Db      469 GAGACTTGGCCCTATCTCAAGTTCGCTCATTTCCCGCTAACCAAGCTATTTCGAAAGCT 528
Qy      303 PheGluGlyLysLysArgValHisValIleAspPheSerMetAenGlnGlyLeuGlnTrp 322
      ::::|||||
Db      529 TTTGAGGGCAAAAAGAGTCCATGTTATCGATTTCTCAATGACCAAGCATGCAATGG 588
Qy      323 ProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArgLeuThr 342
      ::::|||||
Db      589 CTGCTTTGATGCAAGCTTTAGCGTTAAGAGTTGCTGCTCGCGGCTTTTAGGTTAACT 648
Qy      343 GlyIleGlyProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeu 362
      ::::|||||
Db      649 GGGATCGGACCTCTGACCACTCACTCATCTTCAGGAGATGTTGGTGGAGATTG 708
Qy      363 AlaGlnLeuAlaGluAlaIleHisValGluPheGluTyArgGlyPheValAlaAsnSer 382
      ::::|||||
Db      709 GCTCAGTTTCCGAANAGATTCAGTGGAGTTCCAGTATCGTGGTTCTGGCTTAATTCG 768
Qy      383 LeuAlaAspLeuAspAlaSerMetLeuLuleuArgProSerGluThrGluAlaValAla 402
      ::::|||||
Db      769 TTGGCCGATTTGGATGCTTCAATGCTTGATCTCAGGCCGAGTGAAGTTGAAGCCGTGCC 828
Qy      403 ValAenSerValPheGluLeuHisLysLeuGlyArg 415
      ::::|||||
Db      829 GTTAACTCGGTTTCGAGTTACATAAATTTGTGGCTCGG 867

RESULT 13
LOCUS CO124730 845 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_Eb07N03.f GR_Eb Gossypium raimondii cDNA clone GR_Eb07N03 5',
mRNA sequence.
ACCESSION CO124730
VERSION CO124730.1 GI:48823417
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 845)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,

```



```

Db      469  ATTCTCGAAGCTTTTCGCCGTCGCCGAAGAGTTCATGTTATAGATTTTGGTCTTAAACAA 410
Qy      319  GlyLeuGlnTTPProAlaLeuMetClnAlaLeuAlaLeuArgGluGlyGlyProProSer 338
Db      409  GGGATGCAATGGCTTCGCTTATGCAAGCGTGGCAATTACGCTCTGGCGGTCTCTCTACG 350
Qy      339  PheArgLeuThrGlyIleGlyProAlaAlaAspAsnSerAspHisLeuHisGluVal 358
Db      349  TTCGGTTAACCGGCATCGGACCGCGCGGAGGACATATCTGATGCTTTCACACNAGTT 290
Qy      359  GlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyrArgGlyPhe 378
Db      289  GGCTGGAATTTAGCTCAGCTTCGCTCAGACAATCGGTGTTTCAGTTCGAATTCGTGGATT 230
Qy      379  ValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgProSerGluThr 398
Db      229  GTTTGTAAACAGTATTCGGGATCTTGACCCGAATATGCTTGAGATCCGACCCCGGT 176
Qy      399  GluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyValGlyThrGlyGly 418
Db      175  GAAGCTGTTGCTGTTAACTTCGCTTTCGAGCTTCATACCAATGTTAGCTCGACCCGGTTCA 116
Qy      419  IleGluLysValPheGlyValValLysGlnIleLysProValIlePheThrValValGlu 438
Db      115  GTTGAGAAAGTTCTCAACACTGTTTAAAGAAGATAAACCCCTAAAAATCGTAACAATCGTTGAG 56
Qy      439  GlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSerLeu 456
Db      55  CAAGAAGCAAAATCATACCGACCGCGGTTTTCGTAGACCGGTTTTCACCGAAGCGTT 2

```

Search completed: November 3, 2004, 11:22:17
 Job time : 6490 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 2, 2004, 06:34:44 ; Search time 966 Seconds
(without alignments)
3108.353 Million cell updates/sec

Title: US-10-030-194A-4

Perfect score: 2950

Sequence: 1 MKRDLHOFQGNHGTSIAGS.....LSWHTRLITTSAMKLSAVH 572

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO spoal_p/US10030194/runat_01112004.184751.28627/app_query.fasta_1.711
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITIS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10030194 @CGN 1 1.885 @runat_01112004.184751.28627 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	2950	100.0	1779	4 AAF25481	Aaf25481 Nucleotid
2	2946	99.9	1779	4 AAF25480	Aaf25480 Nucleotid
3	2408.5	81.6	1764	4 AAD05776	Aad05776 Arabidops
4	2408.5	81.6	1764	5 AAD06661	Aad06661 A. thalia
5	2408.5	81.6	1764	12 ADO01802	Ado01802 Thalecres
6	2407	81.6	1764	6 AB214708	Ab214708 Arabidops

7	2111	71.6	1602	3 AAC45745	Aac45745 Arabidops
8	2111	71.6	1602	6 AB213221	Ab213221 Arabidops
9	2110.5	71.5	1964	2 AAT91937	Aat91937 Arabidops
10	2106.5	71.4	1951	4 AAD05791	Aad05791 Arabidops
11	2106.5	71.4	1951	5 AAD06646	Aad06646 A. thalia
12	2106.5	71.4	1951	10 ADD55687	Add55687 Thalecres
13	2106.5	71.4	1951	10 ADD30751	Add30751 Plant yie
14	2106.5	71.4	1951	12 ADI43826	Adi43826 Plant tra
15	2106.5	71.4	1951	12 ADO01804	Ado01804 Thalecres
16	1943	65.9	1643	2 AAT91938	Aat91938 Arabidops
17	1929	65.2	1642	2 AAT91941	Aat91941 Arabidops
18	1923	65.2	1642	2 AAT91939	Aat91939 Arabidops
19	1905	64.6	1636	2 AAT91940	Aat91940 Arabidops
20	1602	54.3	2255	2 AAX36280	Aax36280 Maize lai
21	1594	54.0	2385	12 ADM47783	Adm47783 Polynucle
22	1592	54.0	2500	12 ADM47776	Adm47776 Polynucle
23	1568	53.2	2125	2 AAX36279	Aax36279 Wheat Rht
24	1551.5	52.6	2302	12 ADL33445	Adl33445 Festuca a
25	1535.5	52.1	1536	12 ADO61600	Ado61600 Transcrip
26	1390.5	47.1	2709	2 AAX36275	Aax36275 Composite
27	1325	44.9	1768	2 AAX36278	Aax36278 Wheat Rht
28	1205	40.8	1746	2 AAX36285	Aax36285 Consensus
29	940	31.9	1473	12 ADL33433	Adl33433 Lolium pe
30	857	29.1	522	10 ABX57643	Abx57643 Arabidops
31	602	20.4	2032	12 ADI42786	Adi42786 Plant tra
32	597.5	20.3	2537	12 ADM47782	Adm47782 Polynucle
33	587.5	19.9	1473	12 ADO62006	Ado62006 Transcrip
34	586	19.9	1449	6 AB213348	Ab213348 Arabidops
35	586	19.9	1449	10 ADD30187	Add30187 Plant yie
36	586	19.9	1449	10 ADE37170	Ade37170 Plant yie
37	586	19.9	1449	12 ADI41726	Adi41726 Plant tra
38	583.5	19.8	2204	10 ADC47017	Adc47017 Rice CIGR
39	583.5	19.8	2204	10 ADC79241	Adc79241 Rice CIGR
40	581	19.7	1729	2 AAV63099	Aav63099 Tomato ls
41	578.5	19.6	2396	12 ADM47786	Adm47786 Polynucle
42	566	19.2	2163	2 AAT95753	Aat95753 Arabidops
43	564.5	19.1	411	12 ADP92692	Adp92692 Cotton ex
44	555.5	18.8	1634	10 ADD30984	Add30984 Plant yie
45	555.5	18.8	1634	12 ADI44248	Adi44248 Plant tra

ALIGNMENTS

RESULT 1

AAF25481
ID AAF25481 standard; DNA; 1779 BP.
XX
AC AAF25481;
XX
DT 15-MAY-2001 (first entry)
XX
DE Nucleotide sequence of a mutant GRAS protein.
XX
KW GRAS protein; BZH gene; transcription factor; gibberellin; morphogenesis;
KW plant development; dwarf plant; crucifer; ss.
XX
OS Brassica napus.
XX
FH Key Location/Qualifiers
FT CDS 60..1778
FT /*tag= a
FT /product= "GRAS"
XX
XX
PN WO200109356-A1.
XX
PD 08-FEB-2001.
XX
PF 02-AUG-2000; 2000WO-FR002216.
XX
PR 02-AUG-1999; 99FR-00010023.
XX
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA

PI Renard M, Delourme R, Barret P, Brunel D, Froger N, Tanguy X;
XX WPI; 2001-182964/18.
DR P-PSDB; AAB31884.
XX
PT New mutant nucleic acid encoding modified GRAS family protein, used to
PT produce dwarf transgenic plants.
XX
PS Example 1; Page 18-20; 28pp; French.
XX
CC The present sequence encodes a mutant plant protein of the GRAS family.
CC The mutant allele of the BZH gene contains a G1695A mutation resulting in
CC the mutation E546K in the protein. GRAS proteins are transcription
CC factors implicated in regulation of the response to gibberellins and thus
CC in control of morphogenesis and plant development. The mutant GRAS
CC protein is used to produce dwarf plants, specifically crucifers. Dwarf
CC plants may be sown earlier (increasing nitrate accumulation without
CC risking excessive stem growth during winter), and have better resistance
CC to cold and lodging. They are also easier to harvest and allow for better
CC monitoring of the crop
XX
SQ Sequence 1779 BP; 419 A; 430 C; 475 G; 455 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.92e-257 Length: 1779
Score: 2950.00 Matches: 572
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-4 (1-572) x AAF25481 (1-1779)

QY 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
DB 60 ATGAAGAGGGATCTTCATPCAGTTCCAGGTCCCAACCAACCGGACATCAATCGCCGGTTCT 119

QY 21 SerThrSerProAlaValPheGlyLysAspLysMetMetValLysGluGluGlu 40
DB 120 TCCACTTCTTCCCTGCGGTGTTGGTAAACAAGATGATGATGTTCCAAAGAGAAGAA 179

QY 41 AspAspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVal 60
DB 180 GACGACGAGCTTCTAGGAGTCTTGGGTTACAAGGTTAGTCTTCGAGATGGCTGAGGTT 239

QY 61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
DB 240 GCGTTGAACCTCGAGCAGCTTGAGACGATGATGGGTAACGCTCAAGAAAGACGGTTAGCT 299

QY 81 HisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTyrLeuAspAsn 100
DB 300 CACCTCGGAGGATGATCTGTTTCATTAACCCCGCTGAGCTTTACTCGTGGCTTGATTAAC 359

QY 101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
DB 360 ATGCTCAGGAGCTTAACCCACCCGCTCAACGACCGGATCTAACGCTTTGAACCCGAG 419

QY 121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyLysAspLeuLysAlaIleProGly 140
DB 420 ATTAATAATAATAATAACTCGTTTTTTCACCGGAGCGACCTCAAGCGATTCCTCGA 479

QY 141 AsnAlaValCysAtgSerAsnGlnPheAlaPheAlaValAspSerSerSerAsnLys 160
DB 480 AACCGGTTTGTCCAGCATCTAATCAGTTCCGTTTCGCGTTGATTCGTCGANGTAATAG 539

QY 161 ArgLeuLysProSerSerProAspSerMetValThrSerProSerProAlaGlyVal 180
DB 540 CGTTTGAACCGTCTCGAGCCCTGATTCGATGGTTATCTCCATCACCTGCTGGAGTT 599

QY 181 IleGlyThrThrValThrValThrGluSerThrArgProLeuIleLeuValAspSer 200
DB 600 ATAGGAACGAGCGTTTACAACCGGTGACCGAGTCAACTCGTCTTAACTCTCGTGCATCG 659

QY 201 GlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSer 220
DB 660 CAGGACAACGAGTGGCTCTAGTCCACGCGCTTATGGCTGCGCTGAAGCCGTGACAGAGC 719

QY 221 SerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPheLeuAlaValSerGln 240
DB 720 AGCAACTTGACTCTAGCGGAGGCTCTCGTTAAGCAGATTGGTTTCTTGGCCGCTCTCAA 779

QY 241 AlaGlyAlaMetArgLysValAlaThrPheAlaGluAlaLeuAlaAtgArgIleTyr 260
DB 780 GCCGAGCCATGAGAAAGTCGCCACGTACTTCCCGAAGCTCTCCGCGGAGGATCTAC 839

QY 261 ArgLeuSerProProGlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHis 280
DB 840 GCGCTCTCTCGCCGACGACGACATCGATCACTCTTTATCCGATACTCTCCAGATGCAC 899

QY 281 PheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeu 300
DB 900 TTCTACGAGACTTGGCTTACCTCAAGTTCGCTCACTTCACGGGGAATCAGCGGATTCCTC 959

QY 301 GluAlaPheGluGlyLysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeu 320
DB 960 GAGGCTTTCGAAGGGAAGAGAGTCCAGTCACTCATCTTCTCGATGAACCAAGGGCTT 1019

QY 321 GlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArg 340
DB 1020 CAGTGGCCCGGCTTATGCAAGCCCTTGGCTTGAGGGAAGGAGTCTCTCCGAGTTTCAGG 1079

QY 341 LeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCys 360
DB 1080 TTAACCGGAATTGGTCTCTCCGCGCGGATAACTCCGATCATCTCCATGAAGTTGGATGT 1139

QY 361 LysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAla 380
DB 1140 AAGTTGCTCAGCTCGCGAGGCGATTCACTCGAGTTTGAGTATCGTGGCTTTGTGCT 1199

QY 381 AsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAla 400
DB 1200 AATAGCTTAGCTGATCTTGATGCTCGATGCTTCCAGCTTAGACCGAGTGAACCGAAGCT 1259

QY 401 ValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArgThrGlyGlyIleGlu 420
DB 1260 GTGCGGTTAACTCTGTTTTCGAGCTCCACAAGTCTCTAGGCCCTACCGGTGGATAGAG 1319

QY 421 LysValPheGlyValValLysGlnIleLysProValIlePheThrValValGluGlnGlu 440
DB 1320 AAAGTCTTCGCGGTGTGAACAACATTAACCGGTGATTTTCACGGTTGTTGAGCAAGAA 1379

QY 441 SerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSerLeuHisTyrTrpSer 460
DB 1380 TCGAATCATAAACGCTCCGGTTTTCTTAGACCGGTTTACTGAATCGCTGCATTATTATTCG 1439

QY 461 ThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLysValMetSerGluVal 480
DB 1440 ACGTTGTTGATTCCTTGAAGGTGCTCCGAGTAGCCAAGATAAAGTTATGTCGGAAGTT 1499

QY 481 TyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyProAspArgValGluArg 500
DB 1500 TATTTAGGGAACAGATTGCAATCTGTGGCTTCGAGAGGTCCGACCGTGTGAGAGA 1559

QY 501 HisGluThrLeuSerGlnTrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHis 520
DB 1560 CATGAGACGCTGAGTCAATGTCGAACCGGTTTCGTTTCGTCGGTTTTTTCGCGCGGCGAT 1619

QY 521 LeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGlu 540
DB 1620 CTCGGGCTAACCGGTTTAAAGCAAGCAGTACGCTTTTGGCTTTGTTTAAATGGAGGCCAA 1679

QY 541 GlyTyrArgValGluLysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeu 560
DB 1680 GGTATCGTGGAGAGAGATAATGCGGTGTTGATGTTGATGTTGGCACACTCGACCGCTC 1739

QY 561 IleThrThrSerAlaTrpLysLeuSerAlaValHis 572

```
Db      1740 ATAAACCACCTCGCTTGAAGCTCTCGCGGTGCAC 1775
RESULT 2
AAF25480
ID      AAF25480 standard; DNA; 1779 BP.
XX
AC      AAF25480;
XX
DT      15-MAY-2001 (first entry)
XX
DE      Nucleotide sequence of a wildtype GRAS protein.
XX
KW      GRAS protein; BZH gene; transcription factor; gibberellin; morphogenesis;
KW      plant development; dwarf plant; crucifer; ss.
XX
OS      Brassica napus.
XX
FH      Key Location/Qualifiers
FT      CDS 60..1778
FT      /*tag= a
FT      /product= "GRAS"
XX
PN      WO200109355-A1.
XX
PD      08-FEB-2001.
XX
PF      02-AUG-2000; 2000WO-FR002216.
XX
PR      02-AUG-1999; 99FR-00010023.
XX
PA      (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
PI      Renard M, Delourme R, Barret P, Brunel D, Froger N, Tanguy X;
XX
DR      WPI; 2001-182964/18.
DR      P-ESDB; AAB31883.
XX
PT      New mutant nucleic acid encoding modified GRAS family protein, used to
PT      produce dwarf transgenic plants.
XX
PS      Example 1; Page 13-15; 28pp; French.
XX
CC      The present sequence encodes a wild type plant protein of the GRAS
CC      family. The specification describes a mutant allele of the BZH gene,
CC      which contains a G1695A mutation resulting in the mutation E546K in the
CC      protein. GRAS proteins are transcription factors implicated in regulation
CC      of the response to gibberellins and thus in control of morphogenesis and
CC      plant development. The mutant GRAS protein is used to produce dwarf
CC      plants, specifically crucifers. Dwarf plants may be sown earlier
CC      (increasing nitrate accumulation without risking excessive stem growth
CC      during winter), and have better resistance to cold and lodging. They are
CC      also easier to harvest and allow for better monitoring of the crop
XX
SQ      Sequence 1779 BP; 418 A; 430 C; 476 G; 455 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,13e-256 Length: 1779
Score: 2946.00 Matches: 571
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.83% Mismatches: 0
Query Match: 99.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-4 (1-572) x AAF25480 (1-1779)
Qy      1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db      60 ATGAGAGGGATCTTCATCAGTTCCCAAGTCCCAACCCAGCGGACATCAATCGCCGGTTCT 119
Qy      21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
Db      120 TCCACTTCTCCCTCGCGGTGTTTGGTAAAGACAAGATGATGATGGTCAAAAGAAAGAA 179
```

```
Qy      41 AspAspGluLeuLeuGlyValLeuGlyTyrIlyValArgSerSerGluMetAlaGluVal 60
Db      180 GACGACGAGCTTCTAGGAGTCTTGGGTTCACAGGTTTAGGTCTTCGGAGATGGCTGAGGTT 239
Qy      61 AlaLeuLysLeuGluGlnLeuLeuThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db      240 GCGTTGAACCTCGAGCAGCTTGAGCAGATGATGGGTAAACGCTCAAGAAAGACGGTTAGCT 299
Qy      81 HisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTyrLeuAspAsn 100
Db      300 CACCTCGGACGGATACTGTTCAATCAACCCCGCTGAGCTTTACTCGTGGCTTGATAAC 359
Qy      101 MetLeuThrGluLeuAsnProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db      360 ATGCTCAGGAGCTTAAACCCACCGCTGCAACGACCGGATCTAACGCTTTCAACCCGAG 419
Qy      121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
Db      420 ATTAATTAATTAATTAATTAATCTCGTTTTTACCAGGAGCGGACCTCAAAAGCGATTCCTGA 479
Qy      141 AsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerSerAsnLys 160
Db      480 AACCGGTTTGTTCGACATCTAATCAGTTGCGGTTTGGGTTGATTGTCGAGTAATAAG 539
Qy      161 ArgLeuLysProSerSerSerProAspSerMetValThrSerProSerProAlaGlyVal 180
Db      540 CGTTTGAACCGCTCTCGAGCCCTGATTTCGATGCTTACATCTCCATCACCTGCTGGAGTT 599
Qy      181 IleGlyThrThrValThrValThrValThrGluSerThrArgProLeuIleLeuValAspSer 200
Db      600 ATAGGAACGACGGTTTCAACCGTGACCGAGTCAACTCGTCTCTTAATCTCGTGCAGCTCG 659
Qy      201 GlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSer 220
Db      660 CAGGACAAACGAGTGGCTTAGTCCACGCGCTTATGGCTGCGCTGAAGCCGTGACAGAGC 719
Qy      221 SerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPheLeuAlaValSerGln 240
Db      720 AGCAACTGACTCTAGCGGAGGCTCTCGTTAAGCAGATTGGTTCTTCTGGCGGCTCTCTCA 779
Qy      241 AlaGlyAlaMetArgLysValAlaAlaThrPheAlaGluAlaLeuAlaAspArgIleTyr 260
Db      780 GCCGGAGCCATGAGAAAGTCGCCACGTACTTCGCCGAAGCTCTCGCCGCGGAGGATCTAC 839
Qy      261 ArgLeuSerProProGlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHis 280
Db      840 CGCTCTCTCGCGCGCAGACGACATCGATCACTCTTTATCCGATACTCTCCAGATGCAC 899
Qy      281 PheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeu 300
Db      900 TTCTACGAGACTTGCCCTTACCTCAAGTTCCGCTCACTTCAACGGGAAATCAGGCGGATTC 959
Qy      301 GluAlaPheGluGlyLysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeu 320
Db      960 GAGCTTTTCGAAGGAAAGAGAGATCCACGTATCGATTTCTCGATGAACCAAGGGCTT 1019
Qy      321 GlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArg 340
Db      1020 CAGTGGCCCGGCTTATGCAAGCCCTTGGCTTGAGGAGAGAGGAGGCTCTCCGAGTTTCAGG 1079
Qy      341 LeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCys 360
Db      1080 TTAACCGGAATTGGTCTCTCCCGCGCGGATAACTCCGATCATCTCCATGAAGTTGGATGT 1139
Qy      361 LysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAla 380
Db      1140 AAGTTGGCTCAGCTCGCGAGCGGATTCACGTCGAGTTGAGTATCGTGGCTTTGTTGCT 1199
Qy      381 AsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAla 400
Db      1200 AATAGCTTAGCTGATCTTTGATGCTTCGATGCTTGGCTTAGCTTAGCCGAGTGAACCGAAGCT 1259
```

```
QY 401 ValAlaValAanSerValPheGluLeuHisLysLeuLeuGlyArgThrGlyGlyLeuGlu 420
Db 1260 GTGGCGGTAACTCTGTTTTGAGCTCCACAAGCTCTTAGCGGTACCGTGGGATAG 1319
QY 421 LysValPheGlyValValGlyGlnIleLysProValIlePheThrValValGluGlnGlu 440
Db 1320 AAAGTCTTGGCGGTGTGAAAACAGATTAAACCGGTGATTTTTACCGGTTGTTGAGCAAGAA 1379
QY 441 SerAenHisAenGlyProValPheLeuAenArgPheThrGluSerLeuHisTyrTyrSer 460
Db 1380 TCGAATCATACGGTCCGGTTTTCTTAGACCGGTTACTGAATCGCTGCATATATTTCG 1439
QY 461 ThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLysValMetSerGluVal 480
Db 1440 ACGTTGTTTGAATTCCTTGGAGGTGCTCCGAGTAGCCAAAGATAAAGTTATGTGCGAAGTT 1499
QY 481 TyrLeuGlyLysGlnIleCysAenLeuValAlaCysGluGlyProAspArgValGluArg 500
Db 1500 TATTTAGGGAACAGATTTGCAATCTGGTGGCTTCGGAAGGTCCGGACCGGTTTGAGAGA 1559
QY 501 HisGluThrLeuSerGlnTrpSerAenArgPheGlySerSerGlyPheAlaProAlaHis 520
Db 1560 CATGAGACGCTAGTCAATGTGACCGGTTGGTTCGTCGGTTTTGCCCGCGCAT 1619
QY 521 LeuGlySerAenAlaPheLysGlnAlaSerThrLeuLeuAlaLeuPheAenGlyGlyGlu 540
Db 1620 CTCGGTCTAACGGGTTTTAAGCAAGCGAGTAGCGTTTTTGGCTTTTGAATGGAGCGAA 1679
QY 541 GlyTyrArgValGluLysAenGlyCysLeuMetLeuSerTrpHisThrArgProLeu 560
Db 1680 GGTATATCGTGTGGAGAGATAAATGGGTGTTTGATGTTGAGTTGGCACACTCGACCGCTC 1739
QY 561 IleThrThrSerAlaTrpLysLeuSerAlaValHis 572
Db 1740 ATAACCACTCCGCTTGAAGCTCTCGCGGTGCAC 1775

RESULT 3
AAD05776
ID AAD05776 standard; cDNA; 1764 BP.
XX
AC AAD05776;
XX
DT 31-JUL-2001 (first entry)
XX
DE Arabidopsis thaliana transcription factor, G307 cDNA.
XX
KW Transcription factor; biochemical characteristic; controlling element;
KW structural characteristic; developmental characteristic; gene therapy;
KW agricultural biotechnology; plant trait modification; ss.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 1..1764
FT FT /*tag= a
FT FT /product= "Transcription factor, G307"
XX
PN WO200136597-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-US031344.
XX
PR 17-NOV-1999; 99US-0166228P.
XX
PR 17-APR-2000; 2000US-0197899P.
XX
PR 22-AUG-2000; 2000US-0227439P.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (CREE/) CREELMAN R.
PA (YUGG/) YU G.
PA (ADAM/) ADAM L.
PA (RIEC/) RIECHMANN J L.
PA (HEAR/) HEARD J.
```

```
PA (SAMA/) SAMAHA R.
PA (PILG/) PILGRIM M.
PA (PINE/) PINEDA O.
PA (JIAN/) JIANG C.
XX
PI Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R;
PI Pilgrim M, Pineda O, Jiang C;
XX
DR WPI; 2001-335999/35.
DR P-PSDB; AAE01892.
XX
PT Nucleic acids encoding plant transcription factor polypeptides, useful
PT for altering the biochemical characteristics of plants e.g. corn, potato
PT and cotton plants.
XX
PS Claim 4; Page 66-68; 127pp; English.
XX
CC The present sequence is Arabidopsis thaliana transcription factor, G307
CC cDNA. The transcription factor is used for altering a plant's biochemical
CC characteristics. The transcription factor may be used to alter the
CC structure and developmental characteristics of plants such as soybean,
CC wheat, corn, potato, cotton, rice, oilseed rape, sunflower, alfalfa,
CC sugar cane, turf, banana, blackberry, blueberry, strawberry, raspberry,
CC cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes,
CC honey dew, lettuce, mango, melon, onion, papaya, peas, peppers,
CC pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon,
CC rosaceous fruits and/or vegetable brassicas. Transcription factors are
CC key controlling elements of biological pathways and altering expression
CC levels of 1 or more transcription factors can change entire biological
CC pathways in an organism. Therefore manipulating transcription factor
CC levels in plants offers great potential in agricultural biotechnology for
CC modifying a plant's traits. Transcription factor cDNA is useful in gene
CC therapy
XX
SQ Sequence 1764 BP; 431 A; 383 C; 448 G; 502 T; 0 U; 0 Other;

Alignment Scores:
Score: 5.31e-208 Length: 1764
Pred. No.: 2408.50 Matches: 490
Percent Similarity: 85.79% Conservative: 23
Best Local Similarity: 81.94% Mismatches: 48
Query Match: 81.64% Indels: 37
DB: 4 Gaps: 10

US-10-030-194A-4 (1-572) x AAD05776 (1-1764)
QY 1 MetLysArgAspLeuHisGlnPheGlnGly-----ProAenHisGlyThrSerIleAla 18
Db 1 ATGAAGAGAGATCATCCCAATTCCAAGGTTCGATTGTCCAACCCAGCGGACTTCTTTCT 60
QY 19 GlySerSerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGlu 38
Db 61 TCATCATCAATCTCT-----AAAGATAAGATGATGATGATGATGATGATGATGATGAT 102
QY 39 GluGlu-----AspAspGluLeuLeuGlyValLeuGlyTyrLysVal 52
Db 103 GAAGAAGAGCGTGGAGGTTAATCATCGACGACGAGCTTCTCGCTGTTTAGGTACAAAGTT 162
QY 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGly 72
Db 163 AGGTCAATCGGAGATGGCGAGGTTGCTTTGAAATCGAACAATTAGAGACCATGATGATGAT 222
QY 73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92
Db 223 AATGTTCAAGAGAGATGGTTTATCTCATCTCCGACGAGTACTGTTTCATTATTAATCCGTCG 282
QY 93 GluLeuTyrSerTrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
Db 283 GAGCTTTATTTCTTGGCTTGATAATATGCTCTCTGAGCTTAATCTCTCTCTCTCTCTCTCT 342
QY 113 GlySerAenAlaLeuAenProGluIleAenAenAenAenAenAenSerPhe---PheThr 131
Db 343 AGTTCTAACCGGTTTAGATCCGGTTCTTCTTCCGCCGAGAGATTTGTGGTTTTCGCCGCTTCG 402
```

```
QY 132 GlyGlyAspLeuLeuAlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAla 151
DB 403 GATTATGACCTTAAAGTTCATTCGCGAAACGGATT-----TATCAGTTTCGG 450
QY 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
DB 451 ---GGATTTGATCTTCGCTTCCTCGGAATAATCAGAACAGCGTTTGAATTCATGCTCG 507
QY 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182
DB 508 AGTCCTGATCTATGTTATCATCTGACTTCGACGGGTACCGAGATTGGTGGAGTCATAGA 567
QY 183 ThrThrValThrValThr-----GluSerThrArgProLeu 195
DB 568 ACGACGGTGACGACCAACCAACGACGACGCGCGCGGCTGAGTCAACTCGTTCTCTGT 627
QY 196 IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
DB 628 ATCCTGGTTGACTCGCAAGAACCGGTTCGTTTAGTCCACGCGCTTATGCTTTGTGCA 687
QY 216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235
DB 688 GAAGCAATCCACGACAGACAATTTGACTCTAGCGGAAGCTCTTGTCAAGCAATCGATGC 747
QY 236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaIleThrTyrrPheAlaGluAlaLeu 255
DB 748 TTAGCTGTGTCTCAAGCGCGAGCTATGAGAAAGTGGCTACTTACTTCGCGCAAGCTTTA 807
QY 256 AlaArgArgIleTyrrArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAsp 275
DB 808 GCTCGCGCGATCTACCGTCTCTCCGCGCAGAAATCAGATCATGTCTCTCCGAT 867
QY 276 ThrLeuGlnMetHisPheThrCysProTyrrLeuLysPheAlaHisPheThrAla 295
DB 868 ACTCTTCAGATGCATTTACGAGACTTGCTCTTATCTTAATTCGCTCACTTCACGGCG 927
QY 296 AsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgValHisValIleAspPheSer 315
DB 928 AACCAAGCGATTTCTCGAAGCTTTCAAGGTAAGAAGAGAGATACAGCTCATTTGATTC 987
QY 316 MetAsnGlnGlyLeuGlnThrProAlaLeuMetGlnAlaLeuAlaLeuArgGluGly 335
DB 988 ATGAACCAAGGTCITTCATATGGCGCTTCGCTTATGCAAGCTCTTCGCGCTTCGAGAGAG 1047
QY 336 ProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeu 355
DB 1048 CCTCCAACTTTCGGTTAACCGGAATTTGGTCACCGCGCGCGGATAATTTCTGATCATCTT 1107
QY 356 HisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValIleGluPheGluTyrr 375
DB 1108 CATGAAGTTGGTTGTAAATTAGCTCAGCTTCGCGAGCGGATTCACGTTAGAATTCGAATAC 1167
QY 376 ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395
DB 1168 CGTGGATTCGTGTCTAAACAGCTTAGCCGATCTCGATGCTTCGATGCTTGAGCTTAGACCG 1227
QY 396 SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArg 415
DB 1228 AGCGATACGGAAGCTGTTCGGGTGAATCTGTTTGGAGTACATAAGCTCTTAGTGTGT 1287
QY 416 ThrGlyGlyIleGluLysValPheGlyValValLysGlnIleLysProValIlePheThr 435
DB 1288 CCCGTTGGGATAGAAAGTTCTCGCGCTTGTGAACACAGATTAACCGGTGATTTTCACG 1347
QY 436 ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455
DB 1348 GTGGTTGACGAAGAAATCGAACCATTAACGACCGGTTTTCTTAGACCGGTTTTACTGAATCG 1407
QY 456 LeuHisTyrrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys 475
DB 1408 TTACATTATATTACACTCTGTTTATGTTGTTGGAAGGATTCGGAATAGTCAAGACAA 1467
```

```
QY 476 ValMetSerGluValTyrrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro 495
DB 1468 GTCATGCTGAAGTTTACTTAGGGAAACAGATTTTGAATCTGGTGGCTTGTGAAGCTCCT 1527
QY 496 AspArgValGluArgHisGluThrLeuSerGlnTyrSerAsnArgPheGlySerSerGly 515
DB 1528 GACAGATCTGAGAGACACGAAACGTTTGAGTCAATGGGAAACCGGTTTGGTTCGTCGGGT 1587
QY 516 PheAlaProAlaHisLeuLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeu 535
DB 1588 TTAGCGCGCGCACATCTTTGGGTCTAACCGCTTTAAGCAAGCAGATGATGCTTTTGTCTGTG 1647
QY 536 PheAsnGlyGlyGluGlyTyrrArgValGluLysAsnAsnGlyCysLeuMetLeuSerTrp 555
DB 1648 TTTAATAGTGGCCAAAGTTATCGTGTGGAGGAGATTAATGATGTTTATGTTGGTTGG 1707
QY 556 HisThrArgProLeuIleThrSerAlaTrpLysLeuSer---AlaValHis 572
DB 1708 CACACTCGCCCACTCATTTACCACCTCGCTTGGAAACTCTCGACGCGCGCGCAC 1761
RESULT 4
AAD06661
ID AAD06661 standard; cDNA; 1764 BP.
XX
AC AAD06661;
XX
DT 10-AUG-2001 (first entry)
XX
DE A. thaliana transcription factor G308 homolog, G307 cDNA.
XX
KW Plant transcription factor; phenotype; sugar sensing characteristic;
KW transgenic plant; plant yield; growth; germination; photosynthesis;
KW glyoxylate metabolism; respiration; pathogen response; wounding response;
KW cell cycle regulation; pigmentation; flowering; senescence; physiology;
KW storage organ; metabolism; ss.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
CDS 1..1764
FT /*tag= a
FT /product= "Transcription factor homolog"
XX
PN WO200135725-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-US031414.
XX
PR 17-NOV-1999; 99US-0166228P.
PR 17-APR-2000; 2000US-0197899P.
PR 22-AUG-2000; 2000US-0227439P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J.
PA (PINE/) PINEDA O.
PA (PILG/) PILGRIM M.
PA (ADAM/) ADAM L.
PA (RIEC/) RIECHMANN J L.
PA (YUGG/) YU G.
PA (SAMA/) SAMAHA R.
XX
PI Jiang C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann Jr.;
PI Yu G, Samaha R;
XX
XX WPI; 2001-335977/35.
DR P-PSDB; AAE02560.
XX
XX Nucleic acids encoding plant transcription factor polypeptides, useful
XX for altering the sugar sensing characteristics of plants and increasing
XX yield, e.g. corn, potato and cotton plants.
XX
```

Claim 4; Page 113-115; 151pp; English.

CC The patent relates to polynucleotides encoding 35 plant transcription
 CC factors which may be used to modify phenotype associated with a plant's
 CC sugar sensing characteristics and increasing yield when their expression
 CC level is altered. Sugars are central regulatory molecules that control
 CC aspects of physiology, metabolism and development. Therefore the cDNAs
 CC and proteins of the invention are useful for modifying the growth and
 CC germination rates of plants, photosynthesis, glyoxylate metabolism,
 CC respiration, starch and sucrose synthesis and degradation, pathogen
 CC response, wounding response, cell cycle regulation, pigmentation,
 CC flowering and senescence of plants and for modifying sink-source
 CC relationships in seeds, tubers, roots, and other storage organs leading
 CC to an increase in yield. The transcription factor polynucleotides and
 CC polypeptides may be used to alter the structure and developmental
 CC characteristics of plants such as soybean, wheat, corn, potato, cotton,
 CC rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry,
 CC blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower,
 CC coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon,
 CC onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn,
 CC tobacco, tomato, watermelon, rosaceous fruits and/or vegetable brassicas.
 CC The present sequence is a homolog of Arabidopsis thaliana transcription
 CC factor cDNA

SQ Sequence 1764 BP; 431 A; 383 C; 448 G; 502 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 5.31e-208 Length: 1764
 Score: 2408.50 Matches: 490
 Percent Similarity: 85.79% Conservative: 23
 Best Local Similarity: 81.94% Mismatches: 48
 Query Match: 81.64% Indels: 37
 DB: 5 Gaps: 10

US-10-030-194a-4 (1-572) x AAD06661 (1-1764)

QY 1 MetLysArgLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerIleAla 18
 DB 1 ATGAGAGAGATCATCCACCAATTCGAAGTCCATGTCGATGTCACACCGGACTCTCTCTCT 60
 QY 19 GlySerSerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGlu 38
 DB 61 TCATCATCAATCTCT-----AAAGATAAGATGATGTCGAAAAA 102
 QY 39 GluGlu-----AspAspGluLeuGlyValLeuGlyTyrLysVal 52
 DB 103 GAAGAAGACGGTGGAGGTAAACATGACGACGACGAGCTTCGCTGTTTATAGTTTACAAAGTT 162
 QY 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGly 72
 DB 163 AGGTATCGGAGATCGCGGAGGTTCCTTTGAACACTCGAACATTAGACGATGATGATG 222
 QY 73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92
 DB 223 AATGTTCAAGAAGATGGTTTATCTCATCTCGCGACGGATCTGTTCAATTATAATCCGTCG 282
 QY 93 GluLeuTyrSerTyrLeuAsnAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
 DB 283 GAGCTTTATCTTGGCTTGATAATATGCTCTCGAGCTTAATCTCCTCTCTCTCTCTCTCGCG 342
 QY 113 GlySerAsnAlaLeuAsnProGluIleAsnAsnAsnAsnSerPhe---PheThr 131
 DB 343 AGTTCTAACGGTTTAGATCCGGTCTCTCTTCGCGGAGATTTGGGTTTTCGCGCTTCG 402
 QY 132 GlyGlyAspLeuLysAlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAla 151
 DB 403 GATTATGACCTTAAAGTCATCTCCGGAAACCGGAT-----TATCAGTTTCCG 450
 QY 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
 DB 451 ---GCGATTGATCTCTCGTCTTCGCGAATAATCAGAACACGCGTTTGAATCATGCTCG 507
 QY 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182

DB 508 AGTCCTGATTCATGGTTTACATCGACTTCGACGGGTACGCAGATTGGTGGAGTTCATGGA 567
 QY ThrThrValThrThrValThr-----GluSerThrArgProLeu 195
 DB 568 ACGACGGTGCAGCAACACCCACGACGACGACGACGACGACGACGACGACGACGACGACG 627
 QY IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
 DB 628 ATCTGTTGATCTGCAAGAGAACGGTGTGTTAGTCCACGCGCTTATGGCTTGTGCA 687
 QY GluAlaValGlnSerAsnLeuThrIleAlaGluAlaLeuValLysGlnIleGlyPhe 235
 DB 688 GAAGCAATCCAGCAGACAATTTGACTCTAGCGGAAGCTCTGTGAAGCAATCCGATGC 747
 QY LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrThrPheAlaGluAlaLeu 255
 DB 748 TTACTGTGTCTCAAGCGCGAGCTATGAGAAAGTGGCTACTTACTTCGCGGAAGCTTTA 807
 QY AlaArgArgIleTyrArgLeuSerProProGlnThrGlnIleAspHisSerLeuSerAsp 275
 DB 808 GCTCGGGGATCTACCGTCTCTCTCCGCGCAGAAATCAGATCGATCATTTGCTCTCCGAT 867
 QY ThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295
 DB 868 ACTCTTCAGATGCACCTTTACGAGACTTGTCTTATCTTAAATTCGCTCACTTCACGGCG 927
 QY AsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgValHisValIleAspPheSer 315
 DB 928 AACCAAGCGATTCGGAAGCTTTTGAAGGTGAAGAGAGATACACGTCATTTGATTTTCG 987
 QY MetAsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGly 335
 DB 988 ATGAACCAAGTCTTCAATGGCTTCGCTTATGCAAGCTCTTGGCTTCGAGAGAGGAGT 1047
 QY ProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeu 355
 DB 1048 CTTCCAACTTTCCGGTTAAACCGGAATTTGGTCCACCGCGCGGATAATTTCTGATCATCTT 1107
 QY HisGluValGlyCysLysLeuAlaGluAlaGluAlaIleHisValGluPheGluTyr 375
 DB 1108 CATGAAGTTGGTTGTAAATTAGCTCAGCTTCGCGAGGCGATTCACGATAGAAATTCGAATAC 1167
 QY ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395
 DB 1168 CGTGGATTCGTTGCTTAACAGCTTAGCCGATCTCGATGCTTCGATGCTTAGACCTTAGACCG 1227
 QY SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArg 415
 DB 1228 AGCGATACGGAAGCTGTTCGGTGAACCTCTGTTTTTTCAGCTTACATAAGCTCTTAGGTCGT 1287
 QY ThrGlyGlyIleGluLysValPheGlyValLysGlnIleLysProValIlePheThr 435
 DB 1288 CCCGGTGGGATAGAAAAGTTCTCGGGCTGTGAAAACAGATTAACCCGGTATTTTCACG 1347
 QY ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455
 DB 1348 GTGTTTGAAGCAAGATTCGAACCTAAACGACCGGTTTTCTTAGACCGGTTTACTGAAATCG 1407
 QY LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys 475
 DB 1408 TTACATATTATTACGACTCTCTGTTGATTCGTTGAAGGAGTTCGGAATAGTCAAGACAAA 1467
 QY ValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro 495
 DB 1468 GTCATGCTGAAGTTTACTTAGGGAAAACAGATTTGTAACTTGGTGGCTTTGTAAGGTCCT 1527
 QY AspArgValGluArgHisGluThrLeuSerGlnTyrSerAsnArgPheGlySerSerGly 515
 DB 1528 GACAGATTCGAGACACAGAACCGTTGAGTCAATGGGGAACCGGTTTGGTTCGTCGGT 1587
 QY PheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeu 535

Db 1588 TTAGCGCGGCACATCTTGGGCTCAACCGGTTTAAGCAAGCAGTATGCTTTTGTCTGTG 1647
Qy 536 PheAsnGlyGlygluGlyTyArgValGluLysAsnAsnGlyCysLeuMetLeuSerTrp 555
Db 1648 TTTAATAGTGGCCAAAGGTTATCGTGTGAGGAGAGACTAATGATGTTTCATGTTGGTTGG 1707
Qy 556 HisThrArgProLeuLeuIleThrThrSerAlaTrpLysLeuSer---AlaValHis 572
Db 1708 CACACTCGCCCACTCATTTACCACCTCCGCTTGGAACCTCTCGACGGCGCGCAC 1761
RESULT 5
ADO01802
ID ADO01802 standard; cDNA; 1764 BP.
XX AC
XX ADO01802;
XX
XX 01-JUL-2004 (first entry)
XX
XX Thalecress transcription factor cDNA #108.
XX
XX Thalecress; transcription factor; ss; gene; plant; transgenic;
KW abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;
KW phosphate limitation; potassium limitation; nitrogen limitation;
KW hormone sensitivity; disease resistance; sugar sensing; seed germination;
KW flowering; inflorescence architectural change;
KW meristem cell differentiation; phyllotaxy; apical dominance;
KW trichome development; seed development; premature senescence;
KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
KW seed morphology; secondary metabolism; light response; shade avoidance.
XX
OS Arabidopsis thaliana.
XX
XX US2004045049-A1.
XX
XX 04-MAR-2004.
XX
XX 10-APR-2003; 2003US-00412699.
XX
XX 13-SEP-1999; 99US-00394519.
PR 21-JAN-2000; 2000US-00489376.
PR 17-FEB-2000; 2000US-00506720.
PR 22-MAR-2000; 2000US-00532591.
PR 22-MAR-2000; 2000US-00533029.
PR 22-MAR-2000; 2000US-00533030.
PR 22-MAR-2000; 2000US-00533392.
PR 22-MAR-2000; 2000US-00533648.
PR 06-APR-2000; 2000WO-US009448.
PR 16-NOV-2000; 2000US-00713994.
PR 27-MAR-2001; 2001US-00819142.
PR 17-APR-2001; 2001US-00837444.
PR 30-JAN-2002; 2002US-00958131.
PR 14-JUN-2002; 2002US-00171468.
PR 09-AUG-2002; 2002US-00225066.
PR 09-AUG-2002; 2002US-00225067.
PR 09-AUG-2002; 2002US-00225068.
PR 17-DEC-2002; 2002US-0434166P.
PR 25-FEB-2003; 2003US-00374780.
XX
XX (ZHAN/) ZHANG J.
PA (FROM/) FROMM M E.
PA (HEAR/) HEARD J E.
PA (RIEC/) RIECHMANN J L.
PA (ADAM/) ADAM L J.
PA (BROU/) BROUN P E.
PA (PINE/) PINEDA O.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIE J S.
PA (YUGG/) YU G.
PA (JIAN/) JIANG C.
PA (SAMA/) SAMAHA R S.
PA (PIUG/) PILGRIM M L.
PA (CREE/) CREELMAN R A.
PA (DUBE/) DUBELL A N.

PA (RATC/) RATCLIFFE O.
PA (KUMI/) KUMIMOTO R.
XX (SHER/) SHERMAN B K.
PI Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;
PI Pineda O, Reuber TL, Keddle JS, Yu G, Jiang C, Samaha RS;
PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;
XX Sherman BK;
DR WPI; 2004-225755/21.
XX P-PSDB; ADO01803.
XX
XX New transgenic plant, useful in developing phenotypes with altered or
XX improved characteristics or traits.
XX
XX Claim 1; SEQ ID NO 215; 213pp; English.
XX
XX The invention relates to a transgenic plant comprises a recombinant
XX polynucleotide having a polynucleotide sequence or its complementary
XX sequence comprising a sequence encoding a polypeptide, that initiates
XX transcription (i.e. a transcription factor) from Arabidopsis, Soybean,
XX Rice, Rape or Corn, comprising any of the sequences appearing as ADO01588
XX -ADO03527 or ADO03530-ADO03559. Also included are using a transgenic
XX plant to grow a progeny plant, an expression cassette comprising a
XX constitutive, inducible or tissue-specific promoter and a recombinant
XX polynucleotide described above), a host cell comprising the expression
XX cassette, producing a modified plant having a modified trait, identifying
XX a factor that is modulated by or interacts with a polypeptide encoded by
XX the polynucleotide sequence and identifying at least one downstream
XX polynucleotide sequence that is subject to a regulatory effect of any of
XX the polypeptides encoded by the polynucleotide described above. The
XX transgenic plant is useful for producing a plant that has an altered
XX trait e.g. an enhanced tolerance to abiotic stress (increased tolerance
XX to chilling, germination in cold conditions, freezing tolerance, tolerance
XX to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
XX salt, tolerance to phosphate limitation, tolerance to potassium
XX limitation, decreased sensitivity to nitrogen limitation), altered
XX hormone sensitivity, reduced sensitivity to abscisic acid, an altered
XX response to ethylene, disease resistance, altered susceptibility to
XX Botrytis, altered susceptibility to Fusarium, altered susceptibility to
XX Erysiphe, altered susceptibility to Pseudomonas syringae, altered
XX susceptibility to Sclerotinia, altered sugar sensing, improved seed
XX germination and seedling vigor, early flowering, late flowering, extended
XX period of flowering, an inflorescence architectural change, a change in
XX stem bifurcations, a lack of a shoot meristem, reduced meristem cell
XX differentiation, altered phyllotaxy, altered branching pattern, reduced
XX apical dominance, reduced trichome density, ectopic trichome development,
XX altered trichome development, altered stem morphology, increased root
XX growth, increased root hairs, altered seed development, altered cell
XX proliferation/cell differentiation, premature senescence, delayed
XX senescence, lethality, increased necrosis, an increase in seedling or
XX plant size, decreased plant size, a change in leaf morphology, increased
XX altered leaf development, increased leaf size and mass, glossy leaves,
XX leaf cell expansion, change in seed morphology, altered seed coloration,
XX increased seed size, decreased seed size, altered seed shape, change in
XX leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid
XX content, increased leaf insoluble sugars, decreased leaf insoluble
XX sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
XX content, an alteration of leaf glucosinolate content, change in seed
XX biochemistry, an increase in seed oil content, decrease in seed oil
XX content, increase in seed fatty acid content, decrease in seed fatty acid
XX content, increase in seed protein content, decrease in seed protein
XX content, alteration in seed prenyl lipid content, increase in seed
XX sterols, upregulation of genes involved in secondary metabolism, increase
XX in root anthocyanins, increase in plant anthocyanins, and alteration in
XX light response or shade avoidance. The present sequence encodes a
XX thalecress transcription factor of the invention.
SQ Sequence 1764 BP; 431 A; 383 C; 448 G; 502 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.31e-208 Length: 1764

Score: 2408.50 Matches: 490

Percent Similarity:	85.79%	Conservative:	23
Best Local Similarity:	81.94%	Mismatches:	48
Query Match:	81.64%	Indels:	37
DB:	12	Gaps:	10
US-10-030-194A-4 (1-572) x ADO01802 (1-1764)			
QY	1	MetLysArgAspLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerIleAla 18	
DB	1	ATGAAGAGAGATCATCCAAATTCGAAGTCCAGTTGTCCTCAACCGGAGCTTCTTCT 60	
QY	19	GlySerSerThrSerProAlaValPheGlyLysAspLysMetMetValLysGlu 38	
DB	61	TCATCATCAATCTCT-----AAAGATAAGATGATGATGCGTGAATAA 102	
QY	39	GluGlu-----AspAspGluLeuLeuGlyValLeuGlyTyrLysVal 52	
DB	103	GAAGAAGACGGTGGAGGTAAACATGACGACGACGAGCTTCTCGCTGTATTAGGTTACAAAGTT 162	
QY	53	ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGly 72	
DB	163	AGGTCATCGGAGATGCGGAGGTTGCTTTGAACCTCGAACAAATTAGACGATGATGAGT 222	
QY	73	AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92	
DB	223	AATGTTCAAGAAGATGTTTATCTCATCTCGCAGCGGATAGTCTTCAATATAATCCGTCG 282	
QY	93	GluLeuTyrSerTrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112	
DB	283	GAGCTTTATCTTGGCTTGATAATATGCTCTCGAGCTTAATCTCTCTCTCTCTCTCTCT 342	
QY	113	GlySerAsnAlaLeuAsnProGluLeuAsnAsnAsnAsnAsnSerPhe---PheThr 131	
DB	343	AGTTCTAACCGTTTATGATCCCGGTTCTTCTTCCCGGAGATTTGGTTTTCGCGCTTCG 402	
QY	132	GlyGlyAspLeuLysAlaIleProGlyAsnAlaValAlaCysArgArgSerAsnGlnPheAla 151	
DB	403	GATTATGACCTTAAAGTCATTTCCCGAACAACGGATT-----TATCAGTTTCCG 450	
QY	152	PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166	
DB	451	---GCGATTGATTTCTCTCTCTCGTGAATAATCAGAACAAAGCGTTTGAATCATGCTCG 507	
QY	167	SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182	
DB	508	AGTCTCTGATTTATGTTTACATCGACTTCGACGGGTACGAGATGCTGGAGTCAATAGGA 567	
QY	183	ThrThrValThrThrValThr-----GluSerThrArgProLeu 195	
DB	568	ACGACGGTGACGACAAACCAACGACGACGACGCGCGGCGTGAAGTCAACTCGTTCTGTT 627	
QY	196	IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215	
DB	628	ATCCTGGTTGATCTCGCAAGAACGGTGTCTGTTTAGTCCACGCGCTTATGGCTTGTGCA 687	
QY	216	GluAlaValGlnSerSerAsnMetThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235	
DB	688	GAAGCAATCCAGACGACAAATTTGACTCTAGCGGAGCTCTTGAAGCAAAATCGATGCG 747	
QY	236	LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrTyrPheAlaGluAlaLeu 255	
DB	748	TTAGCTGTGTCTCAAGCGGAGCTATGAGAAAGTGGCTACTTACTTCGCGGAGCTTTA 807	
QY	256	AlaArgArgIleTyrArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAsp 275	
DB	808	GCTCGCGGATCTACCGTCTCTCTCCGCGGAGAAATCAGATCGATCATTTGCTCTCCCAT 867	
QY	276	ThrLeuGlnMetHisPheThrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295	
DB	868	ACTCTTCAGATGCACCTTTTACGAGACTTGTCTTATCTTAAATTCGCTCCTTACGCGCG 927	
QY	296	AsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgValHisValIleAspPheSer 315	

DB	928	AACCAAGCGATTTCTCGAAGCTTTTGAAGGTAAGAAGAGAGTACACGCTATTGATTCTCG 987	
QY	316	MetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGly 335	
DB	988	ATGAACCAAGGTCTTCAATGGCCCTTATGAAGCTCTTGGCTTCGAGAAGGAGT 1047	
QY	336	ProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeu 355	
DB	1048	CCTCCAACTTCCGGTTAACCGGAATTTGGTCCACCGCGCGGATAAATCTGATCATCTT 1107	
QY	356	HisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyr 375	
DB	1108	CATCAAGTTGGTTCTAAATTAGCTCAGCTTCGCGAGGCGATTACGTAGAAATTCGAATAC 1167	
QY	376	ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395	
DB	1168	CGTGGATTCTGTCTAACAGCTTAGCCGATCTCGATGCTTCGATGCTTAGACCG 1227	
QY	396	SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArg 415	
DB	1228	AGCGATACGGAAGCTGTTGCGGTGAACCTCTGTTTTTCAGCTTACATAAGCTCTTAGTCT 1287	
QY	416	ThrGlyGlyIleGluLysValPheGlyValLysGlnIleLysProValIlePheThr 435	
DB	1288	CCCGTGGGATAGAGAAAGTTCTCGCGGTTGTGAACAGATTAAACCGGTGATTTTCACG 1347	
QY	436	ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455	
DB	1348	GTGTTGAGCAAGAAATCGAACCAACGACCGGTTTTCTTAGACCGGTTTACTGAAATCG 1407	
QY	456	LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys 475	
DB	1408	TTACATATTATTTCGACTCTGTTTGTATTCGTTGAGGAGTTCCGAATAGTCAAGACAAA 1467	
QY	476	ValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro 495	
DB	1468	GTCTGTCTGAAGTTTACTTAGGGAAAACAGATTTGTAATCTGGTGGCTTGTGAAGGTCT 1527	
QY	496	AspArgValGluArgHisGluThrLeuSerGlnTrpSerAsnArgPheGlySerSerGly 515	
DB	1528	GACAGGTGAGAGACACGAAACGTTGAGTCAATGGGAAAACCGGTTTGGTTGCTCGGT 1587	
QY	516	PheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeu 535	
DB	1588	TTAGCGCGCGCACATCTTGGGTCTAACCGGTTTAAGCAAGCGAGTATGCTTTTGTCTGTG 1647	
QY	536	PheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGlyCysLeuMetLeuSerTrp 555	
DB	1648	TTTAATAGTGGCCAAGGTATCGGTGTGAGGAGAGTAAATGGATGTTTGTATGTTGGTTGG 1707	
QY	556	HisThrArgProLeuIleThrThrSerAlaTrpLysLeuSer---AlaValHis 572	
DB	1708	CACACTCGCCCACTCATTTACCACCTTCGCTTGGAAACTCTTCGACGGCGGCGCAC 1761	
RESULT 6			
ID	ABZ14708	standard; DNA; 1764 BP.	
XX	ABZ14708		
AC	ABZ14708;		
XX	21-JAN-2003	(first entry)	
DE	Arabidopsis thaliana	stress regulated gene SEQ ID NO 2513.	
XX	Arabidopsis thaliana;	plant; gene; stress; transgenic; ds.	
OS	Arabidopsis thaliana.		
PN	WO200216655-A2.		
XX	28-FEB-2002.		
PD	24-AUG-2001;	2001WO-US026685.	
PF			

XX 24-AUG-2000; 2000US-0227866P.
 PR 26-JAN-2001; 2001US-0264647P.
 PR 22-JUN-2001; 2001US-0300111P.
 XX
 XX (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Harper JF, Krepes J, Wang X, Zhu T;
 PI WPI; 2002-304127/34.
 XX
 XX Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.
 XX
 PS Claim 144; SEQ ID NO 2513; 577pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 XX
 SQ Sequence 1764 BP; 433 A; 380 C; 449 G; 502 T; 0 U; 0 Other;
 XX

Alignment Scores:
 Pred. No.: Length: 1764
 Score: 2407.00 Matches: 488
 Percent Similarity: 86.03% Conservative: 23
 Best Local Similarity: 82.15% Mismatches: 47
 Query Match: 81.59% Indels: 36
 DB: 6 Gaps: 9

US-10-030-194A-4 (1-572) x ABZ14708 (1-1764)

QY 1 MetLysArgAspLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerIleala 18
 DB 1 ATGAAGAGAGATCATCACCATTCCAAAGTCCAGGTCGATTGTCCACACCGCGGACTTCTTCTTCA 60
 QY 19 GlySerSerThrSerProAlaValPheGlyLysAspLysMetMetValLysGlu 38
 DB 61 TCATCATCAATCTCT-----AAAGATAAGATGATGATGGTGAAGAAA 102
 QY 39 GluGlu-----AspAspGluLeuLeuGlyValLeuGlyTyrLysVal 52
 DB 103 GAAGAAGACGGTGGAGGTAAACATGACGACGAGCTTCTCGTGTGTTTAGGTATCAAGTT 162
 QY 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGly 72
 DB 163 AGGTCAATCGAGATGCGCGAGGTGCTTTGAAACTCGAACAAATTAGACACCATGATGAGT 222
 QY 73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92
 DB 223 AATGTTCAAGAAGATGGTGTATCTCATCTCGCGACGGATACGTGTTTATTAATCCGTCG 282
 QY 93 GluLeuTyrSerTyrLeuAspAsnMetLeuThrGluLeuLeuAsnProProAlaAlaThrThr 112
 DB 283 GAGCTTTATCTTGGCTTGATAATATGCTCTCTGAGCTTAATCTCTCTCTCTCTCTCTCTCCGCG 342
 QY 113 GlySerAsnAlaLeuAsnProGluLeuAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 131
 DB 343 AGTCTCAACGGTTTAGATCCGGTCTCTCTCTCGCGGAGATTTGTGGTCTTTCGCGCTTCG 402
 QY 132 GlyGlyAspLeuLysAlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAla 151
 DB 403 GATTATGACCTTAAAGTCAITCCCGGAAACGCGATT-----TATCAGTTTCCG 450

QY 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
 DB 451 ---GCGATTGATCTTCGTCGAATAATCAGAACAGAGCGTTTGAATCATGCTCG 507
 QY 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182
 DB 508 AGTCTGATTTCTATGTTTACATCGACTTCGACGGGTACGACGATTGGTGGAGTCATAGGA 567
 QY 183 ThrThrValThrThrValThr-----GluSerThrArgProLeu 195
 DB 568 ACGACGGTGACGACAAACACACGACGACGCGCGGGTGAGTCACTGGTCTGTT 627
 QY 196 IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
 DB 628 ATCTCTGGTTGACTCGCAAGAGAACGGTGTCTGTAGTCCACGCGCTTATGGCTGTGCA 687
 QY 216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235
 DB 688 GAAGCAATCCAGCAGAACAAATTTGACTCTAGCGGAAGCTTTGTGAAGCAATCGGATCG 747
 QY 236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrTyrPheAlaGluAlaLeu 255
 DB 748 TTAGCTGTGTTCTCAAGCGCGAGCTATGAGAAAGTGGTACTTACTTCCGCCAAGCTTTA 807
 QY 256 AlaArgArgIleTyrArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAsp 275
 DB 808 GCGCGCGGATCTACCGTCTCTCTCCGCGCAGAAATCAGATCGATCATTTGCTCTCCGAT 867
 QY 276 ThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295
 DB 868 ACTCTTCAGATGCACCTTTTACGAGACTTGTCTTATCTTAAATTCGCTCACTTCACGGCG 927
 QY 296 AsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgValHisValIleAspPheSer 315
 DB 928 AACCAAGCGATTCTCGAAGCTTTTGAAGGTAAAGAGAGATACACGTCATTGATTTCTCG 987
 QY 316 MetAsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLeuAlaLeuArgGluGly 335
 DB 988 ATGAACCAAGGTCTTCAATGCGCTGCACTTATGCAAGCTTTGCGCTTCGACAGAGAGGT 1047
 QY 336 ProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeu 355
 DB 1048 CTTCAACTCTCGGTAAACCGGAATGCTGTCACCGCGCGCGGATAAATTTCTGATCATCTT 1107
 QY 356 HisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyr 375
 DB 1108 CATGAAGTTGGTTGTAAATTAGCTCAGCTTCGGAGGCGATTTCAGTGAATTCGAATAC 1167
 QY 376 ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395
 DB 1168 CGTGGATTTCGTTGCTAACAGCTTAGCCGATCTCGATGCTTCGATGCTTAGAGCTTAGACCG 1227
 QY 396 SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuGlyArg 415
 DB 1228 AGCGATACGGAAGCTGTTGCGGTGAACCTCTGTTTTGAGCTACATAAGCTCTTAGTCTGT 1287
 QY 416 ThrGlyGlyIleGluLysValPheGlyValLysGlnIleLysProValIlePheThr 435
 DB 1288 CCCGTTGGGATAGAGAAGTTCTCGCGGTGTGAAACAGATTAAACCGGTGATTTTCAG 1347
 QY 436 ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455
 DB 1348 GTGGTTGAGCAAGAAATCGAACCAATACGACCGGTTTTCTTAGACCGGTTTACTCAATCG 1407
 QY 456 LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys 475
 DB 1408 TTACATTATTATTCGACTCTCTGTTGATTCGTTGGAAGAGATTCCGAATAGTCAAGACAAA 1467
 QY 476 ValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro 495
 DB 1468 GTCATGCTGAAGTTTACTTAGGGGAAACAGATTTGTAAATCTGGTGGCTTGTGAAGGCTCT 1527

```
Qy 496 AspArgValGluArgHisGluThrIleuSerGlnTrpSerAsnArgPheGlySerSerGly 515
Db 1528 GACAGAGTCGAGAGACACGAAACGTTGAGTCAATGGGAAACCGGTTTGGTTGTCGCGGT 1587
Qy 516 PheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnIlaSerThrLeuLeuAlaLeu 535
Db 1588 TTAGCGCGGCACATCTTGGGTCTAAACGCGTTTAAGCAAGCGAGTATGCTTTTGTCTGTG 1647
Qy 536 PheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGlyCysLeuMetLeuSerTrp 555
Db 1648 TTTTAATAGTGGCCAAAGGTTATCGTGTGGAGGAGATAATGGATGTTTGAATGTTGGTTGG 1707
Qy 556 HisThrArgProLeuIleThrThrSerAlaTrpLysLeuSer 569
Db 1708 CACACTCGTCCACATCAATTACCACTTCGCTTGGAACCTCTCG 1749

RESULT 7
AAC45745
ID AAC45745 standard; DNA; 1602 BP.
XX
AC AAC45745;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47609.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD
PF 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135622P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147018P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
```


QY 137 AlalleProGlyAsnAlaValCysArgSerAsnGlnPheAlaPheAlaValAspSer 156
DB 310 GCTATTCCGGGTGACCGCATTC-----AATCAGTTGCGTATCGATTTCGGCTTCT 360
QY 157 SerSer-----AsnLysArgLeuLysProSer 165
DB 361 TCGTCTAACCAAGCGCGGAGGAGATAGGTATACATAACAACGCGGTGAATGCTCA 420
QY 166 SerSerProAspSerMetValThrSerProSerProAlaGlyValIleGlyThrVal 185
DB 421 AAC-----GGCGTCGTGGAAACCACT--- 441
QY 186 ThrThrValThrGluSerThrArgProLeuLeuValAspSerGlnAspAsnGlyVal 205
DB 442 ACAGCGAGCGGTGAGTCAACTCGGATGTTGCTCTGTTGATCGCAGGAGAACGTTGTG 501
QY 206 ArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeu 225
DB 502 CGTCTCGTTACGCGCTTTGGCTTGGCTGAGCTGTTCAGAAAGAGAATCTGACTGTA 561
QY 226 AlaGluAlaLeuValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArg 245
DB 562 GCGGAAGCTCTGGTGAAGCAATCGGATTCCTAGCGGTTTCTCAATCGAGCGATGAGA 621
QY 246 LysValAlaThrTyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProPro 265
DB 622 AAAGTCGCTACTTACTTCGCGAAGCTCTCGCGCGGAGTTTACCGTCTCTCCGTCG 681
QY 266 GlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCys 285
DB 682 CAGAGTCCAAATCGACCACTCTCTCCGATATCTTCAGATGCACTCTCAGAGACTTGT 741
QY 286 ProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGly 305
DB 742 CCTTATCTCAAGTTCGCTCACTTCACGCGAATCAAGCGATTCGGAAGCTTTTCAAGG 801
QY 306 LysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeu 325
DB 802 AAGAAAGAGTTTCAATGATTTCTATGATCAAGGTCTTCAATGCGCGCGCTT 861
QY 326 MetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGly 345
DB 862 ATGACAGGCTCTTGGCGCTTCGACCTGGTCTCTCTGTTTCCGGTTAACCGGAATTGGT 921
QY 346 ProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeu 365
DB 922 CCACCGGACCGGATAATTCGATTATCTCATGAAGTTGGGTGAAGCTGCCTCATTTA 981
QY 366 AlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAsp 385
DB 982 GCTGAGGCGATTACAGTTGAGTTTGAGTACAGAGGATTTGTGGCTAACACTTTAGCTGAT 1041
QY 386 LeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSer 405
DB 1042 CTTGATGCTTCGATGCTTGAAGCTTGAAGCAAGTGAAGTTGAATCTGTTGCGGTAACTCT 1101
QY 406 ValPheGluLeuHisLysLeuLeuGlyArgThrGlyIleGluLysValPheGlyVal 425
DB 1102 GTTTTCGAGCTTCACAAGCTCTTGGGACGACCTGTCGATCGATAAGTTCTTGGTGTG 1161
QY 426 ValLysGlnIleLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGly 445
DB 1162 GTGAATCAGATTAAACCGGAGATTTTCACTGTGGTTGAGCAGGAATCGAACATAATAGT 1221
QY 446 ProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSer 465
DB 1222 CCGATTTTCTTAGATCGGTGTACTGAGTCGTGTGCAATTATTACTCGACCTGTTGTGACTCG 1281
QY 466 LeuGluGlyAlaProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGln 485
DB 1282 TTGGAGGTGTACCGAGTGGTCAAGACAGAGGTCACTGCGAGGTTTACTTGGGTAAACAG 1341
QY 486 IleCysAsnLeuValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSer 505

DB 1342 ATTCGCAACGTTGTCGCTTGTGATGGACCTCACCAGTTGACCGTCATGAACGTTGAGT 1401
QY 506 GlnTrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAla 525
DB 1402 CAGTGGAGGAACCGGTTCCGGTCTGCTGGTTCGGGCTGCACATATTTGGTTCGAATCGG 1461
QY 526 PheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGlu 545
DB 1462 TTTAAGCAAGCGAGTATGCTTTTGGCTCTGTTCAACGCGGTGAGGGTTATCGGTTGGAG 1521
QY 546 LysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAla 565
DB 1522 GAGAGTGAACGCTGCTCATGTTGGTGTGGCACACACCGCTCATAGCCACCTCGGCT 1581
QY 566 TrpLysLeuSer 569
DB 1582 TGGAACTCTCC 1593
RESULT 9
AAT91937
ID AAT91937 standard; DNA; 1964 BP.
XX
AC AAT91937;
XX
DT 19-MAR-1998 (first entry)
XX
DE Arabidopsis thaliana gibberellin insensitivity gene gai.
XX
KW Gibberellin insensitivity; gai; plant growth inhibition; dwarf phenotype;
lodging resistance; increased yield; flowering regulation;
bolting inhibition; spinach; lettuce; antibody; identification; probe;
primer; antisense; sense; expression regulation; co-suppression; rice;
Bakane disease resistance; ss.
XX
OS Arabidopsis thaliana.
XX
FN WO9729123-A2.
XX
PD 14-AUG-1997.
XX
PF 12-FEB-1997; 97WO-GB000390.
XX
PR 12-FEB-1996; 96GB-00002796.
XX
PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
XX
PI Harberd NP, Peng J, Carol P, Richards DE;
XX
DR WPI; 1997-415295/38.
DR P-PSDB; AAW30792.
XX
PT Nucleic acid encoding gibberellin inhibitor GAI and related antisense
sequences - used to create tall, or particularly, dwarf plants,
especially crops such as maize, rice and wheat.
XX
PS Claim 2; Fig 3; 76pp; English.
XX
CC The present sequence encodes the Arabidopsis thaliana gibberellin
insensitivity (gai) gene product (GAI), the expression of which inhibits
plant growth. However the inhibition is antagonised by gibberellin (GA),
while gai expression confers a dwarf phenotype that is insensitive to GA.
CC Manipulating gai and GAI expression can produce tall or dwarf plants,
particularly the latter for increased resistance to lodging and increased
yield. It may also allow regulation of flowering, i.e. plants remain in
the vegetative state until treated with GA, useful to inhibit bolting in
spinach and lettuce. GAI can be used to raise specific antibodies for
identifying homologous proteins or genes in other species. Gai fragments
can also be used as probes or primers to identify and clone related
sequences, or in the preparation of antisense or sense expression
regulating (co-suppressing) sequences. Rice plants that express GAI may
be resistant to Bakane disease. Manipulation of gai and GAI makes it
possible to tailor the degree of dwarfism and GA sensitivity to

CC particular crops or situations

XX Sequence 1964 BP; 489 A; 426 C; 474 G; 575 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,98e-181 Length: 1964
Score: 2110.50 Matches: 429
Percent Similarity: 79.45% Conservative: 35
Best Local Similarity: 73.46% Mismatches: 51
Query Match: 71.54% Indels: 69
DB: 2 Gaps: 9

US-10-030-194A-4 (1-572) x AAT91937 (1-1964)

QY 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
DB 209 ATGAAGAGAGATCATCAT- : : : : : CATCATCAAGATAAG- 244
QY 21 SerThrSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
DB 245 -----AAGACTATGATGATGAATGAAGAGAC 271
QY 41 Asp-----AspGluLeuGlyValLeuGlyTyrLysValArgSerSerGlu 56
DB 272 GACGGTAACGGCATGGATGAGCTTCTAGCTGTTCTTGGTTACAAGGTTAGGTCATCGGAA 331
QY 57 MetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGlu 76
DB 332 ATGGCTGATGTGCTCAGAAATCGAGCAGCTTGAAGTTATGATGTCTAATGTTCAAGAA 391
QY 77 AspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSer 96
DB 392 GACGATCTTTCTCACTCGCTACTGAGACTGTCTACTATATATCGGCGAGCTTTACACG 451
QY 97 TrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAla 116
DB 452 TGGCTTGATTTCTATGCTCACCGACCTTAATCTCGTCTGCTCT----- 493
QY 117 LeuAsnProGluIleAsnAsnAsnAsnAsnSerPheThrGlyGlyAspLeuLys 136
DB 494 ---AACGCCGAGTAC-----GATCTTAA 514
QY 137 AlaIleProGlyAsnAlaValCysArgSerAsnGlnPheAlaPheAlaValAspSer 156
DB 515 GCTATTCGGGTGACGCGATCTC-----AATCAGTTCTGCTATCGATTCGGCTTCT 565
QY 157 SerSer-----AsnLysArgLeuLysProSer 165
DB 566 TCGTCTAACCAAGCGCGGAGGATACGTATATACTACAAACAAGCGGTTGAAATGCTCA 625
QY 166 SerSerProAspSerMetValThrSerProAlaGlyValIleGlyThrThrVal 185
DB 626 AAC-----GGCGTCGTGGAAACCAACC- 646
QY 186 ThrThrValThrGluSerThrArgProLeuIleLeuValAspSerGlnAsnGlyVal 205
DB 647 ACAGCGCGGTGATCACTCGGATGTTCTGCTGTTGACTCGCAGGAGNACGGTGTG 706
QY 206 ArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeu 225
DB 707 GGTCTCGTTACGGCGCTTTGGCTTGGCTGAGCTGAGCTGTTTCAAGAGGAGATCTGACTGTG 766
QY 226 AlaGluAlaLeuValLysGlnIleGlyPheIleuAlaValSerGlnAlaGlyAlaMetArg 245
DB 767 GCGGAAGCTCTGGTGAAGCAATCGGATCTTAGCTGTTTCTCAATCGGAGCTATGAGA 826
QY 246 LysValAlaThrThrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProPro 265
DB 827 AAAGTCGTACTTACTTCGCGAAGCTCTCGCGCGCGGATTTACCGTCTCTCTCCGTCG 886
QY 266 GlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCys 285
DB 887 CAGAGTCCAATCGACCACTCTCTCCGATACTCTTCAGATGCACTTCTACGAGACTTGT 946

QY 286 ProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGlyGly 305
DB 947 CTTTATCTCAAGTTCGCTCACTTCACGCGGATCAAGGATTCGAGACTTTTCAAGGG 1006
QY 306 LysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeu 325
DB 1007 AAGAAAAGAGTTTCATGTCATTTCTTATGAGTCAAGGTCTTCAATGGCGCGCTT 1066
QY 326 MetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGly 345
DB 1067 ATGCAGGCTCTTGGCTTCGACCTGGTGGTCTCTCTGTTTCCGGTTAAACCGGAATTGGT 1126
QY 346 ProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeu 365
DB 1127 CCACCGGACCGGATTAATTCGATTATCTTCATGAGTTGGGTGTAAGCTGCTCATTTA 1186
QY 366 AlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAsp 385
DB 1187 GCTCAGCGGATTCAGTTGAGTTGAGTACAGAGGATTTGTGGCTAACACACTTAGCTGAT 1246
QY 386 LeuAspAlaSerMetLeuLeuLeuArgProSerGluThrGluAlaValAlaValAsnSer 405
DB 1247 CTTGATGCTTCGATGCTTGAAGCTTAGACCAAGTGAAGTTGAATCTGTTCGGTTAACTCT 1306
QY 406 ValPheGluLeuHisLysLeuLeuGlyArgThrGlyGlyIleGluLysValPheGlyVal 425
DB 1307 GTTTTCAGCTTCAACAGCTCTTGGGACGACCTGGTGGATCGATAGGTTCTTGGTGTG 1366
QY 426 ValLysGlnIleLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGly 445
DB 1367 GTGAATCAGATTTAAACCGGAGATTTTCACTGTGTGTTGAGCAGGAATCGAACCAATAAGT 1426
QY 446 ProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSer 465
DB 1427 CCGATTTTCTTAGATCGGTTTACTGAGTCGTGCAATTATTACTCGACGTTGTTGACTCG 1486
QY 466 LeuGluGlyAlaProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGln 485
DB 1487 TTGAAGGTGTACCGAGTGTGTCAGACAGAGTCAATGTCGGAGGTTTACTTGGGTAAACAG 1546
QY 486 IleCysAsnLeuValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSer 505
DB 1547 ATCTCAACGTTGTGGCTTGTGATGAGACCTGACCGAGTTGAGCGCTCATGAAACGTTGAGT 1606
QY 506 GlnTrpSerAsnArgPheGlySerSerClyPheAlaProAlaHisLeuGlySerAsnAla 525
DB 1607 CAGTGGAGGAACCGGTTGCGGTCTGCTGGGTTCGCGCTGCACATATTTGGTTCGAATCGG 1666
QY 526 PheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGlyTyrArgValGlu 545
DB 1667 TTTAAGCAAGCGAGTATGCTTTTGGCTCTGTTCAACGGCGGTGAGGTTATCGGGTGGAG 1726
QY 546 LysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAla 565
DB 1727 GAGAGTCAAGCGCTGCTCTCATGTTGGGTGGCACACAGACCGCTCATAGCCACCTCGGCT 1786
QY 566 TrpLysLeuSer 569
DB 1787 TGGAAACTCTCC 1798
RESULT 10
AAD05791
ID AAD05791 standard; cDNA; 1951 BP.
XX
AC AAD05791;
XX
DT 31-JUL-2001 (first entry)
XX
DE Arabidopsis thaliana transcription factor, G308 cDNA.
XX
KW Transcription factor; biochemical characteristic; controlling element;
KW structural characteristic; developmental characteristic; gene therapy;


```
Qy 369 eHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSe 389
Db 1185 TCAGTTGAGTTTGAGTACAGAGATTGTGGCTAAACCTTTAGCTGATCTTGATGCTTC 1244
Qy 389 rMetLeuGluLeuArgProSerGluThrGluAlaValAlaAsnSerValPheGluLe 409
Db 1245 GATGCTTGAGCTTAGACCAAGTGAGATTGAATCTGTTGGCGTTAACTCTGTTTCGAGCT 1304
Qy 409 uHisLysLeuLeuGlyArgThrGlyGlyLeGluLysValPheGlyValLysGlnIl 429
Db 1305 TCACAAGCTCTTGGAGCACCTGGTGGCATCGATAAGGTTCTTGCTGTGTGGTGAATCGAT 1364
Qy 429 eLysProValIlePheThrValGluGlnGluSerAsnHisAsnGlyProValPheLe 449
Db 1365 TAAACCGAGATTTTCACTGTGGTTGAGCAGGAATCGAACCATAATAGTCGATTTTCTT 1424
Qy 449 uAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAl 469
Db 1425 AGATCGGTTTACTGAGTCGTTGCATTATTACTCGACGTTGTTTGCATCGTTGGAAGGTGT 1484
Qy 469 aProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLe 489
Db 1485 ACCGAGTGGTCAAGACAAGGTCATGTGCGAGGTTTACTTGGGTAAACACAGATCTGCAAGCT 1544
Qy 489 uValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAs 509
Db 1545 TGTGGCTTGTGATGACCTGACCGAGTTGACGTCATGAACGTTGAGTCAGTGGAGAA 1604
Qy 509 nArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAl 529
Db 1605 CCGGTTCCGGTCTGCTGGGTTTGGCTGCACATATTGTTTGAATGCTTTTAAGCAAGC 1664
Qy 529 aSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGl 549
Db 1665 GAGTATGCTTTTGGTCTGTTTCAACGGCGGTGAGGGTTATCGGGTGGAGGAGTGAACGG 1724
Qy 549 yCysLeuMetLeuSerTrpHisThrArgProLeuIleThrSerAlaTrpLysLeuSe 569
Db 1725 CTGTCTCATGTTGGTTGGCACACACGACCCGCTCATAGCCACCTCGGCTTGGAAACTCTC 1784
Qy 569 r 569
Db 1785 C 1785
RESULT 11
AAD06646
ID AAD06646 standard; cDNA; 1951 BP.
XX
AC AAD06646;
XX
DT 10-AUG-2001 (first entry)
XX
DE A. thaliana transcription factor G308 cDNA.
XX
KW Plant transcription factor; phenotype; sugar sensing characteristic;
KW transgenic plant; plant yield; growth; germination; photosynthesis;
KW glyoxylate metabolism; respiration; pathogen response; wounding response;
KW cell cycle regulation; pigmentation; flowering; senescence; physiology;
KW storage organ; metabolism; ss.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 196..1794
FT FT /*tag= a
FT FT /product= "transcription factor"
XX
PN WO200135725-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-US031414.
```

```
XX 17-NOV-1999; 99US-0166228P.
PR 17-APR-2000; 2000US-0197899P.
PR 22-AUG-2000; 2000US-0227439P.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J.
PA (PINE/) PINEDA O.
PA (PILG/) PILGRIM M.
PA (ADAM/) ADAM L.
PA (RIEC/) RIECHMANN J L.
PA (YUGG/) YU G.
PA (SAMA/) SAMAHA R.
XX
XX Jiang C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann JL;
PI Yu G, Samaha R;
XX
XX WPI; 2001-335977/35.
DR P-PSDB; AAE02545.
XX
XX Nucleic acids encoding plant transcription factor polypeptides, useful
PT for altering the sugar sensing characteristics of plants and increasing
PT yield, e.g. corn, potato and cotton plants.
XX
XX Claim 4; Page 72-74; 151pp; English.
XX
XX The patent relates to polynucleotides encoding 35 plant transcription
CC factors which may be used to modify phenotype associated with a plant's
CC sugar sensing characteristics and increasing yield when their expression
CC level is altered. Sugars are central regulatory molecules that control
CC aspects of physiology, metabolism and development. Therefore the cDNAs
CC and proteins of the invention are useful for modifying the growth and
CC germination rates of plants, photosynthesis, glyoxylate metabolism,
CC respiration, starch and sucrose synthesis and degradation, pathogen
CC response, wounding response, cell cycle regulation, pigmentation,
CC flowering and senescence of plants and for modifying sink-source
CC relationships in seeds, tubers, roots, and other storage organs leading
CC to an increase in yield. The transcription factor polynucleotides and
CC polypeptides may be used to alter the structure and developmental
CC characteristics of plants such as soybean, wheat, corn, potato, cotton,
CC rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry,
CC blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower,
CC coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon,
CC onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn,
CC tobacco, tomato, watermelon, rosaceous fruits and/or vegetable brassicas.
CC The present sequence is an Arabidopsis thaliana transcription factor cDNA
XX
XX SQ Sequence 1951 BP; 482 A; 419 C; 475 G; 575 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 1.37e-180 Length: 1951
Score: 2106.50 Matches: 428
Percent Similarity: 80.55% Conservative: 40
Best Local Similarity: 73.67% Mismatches: 55
Query Match: 71.41% Indels: 58
DB: 5 Gaps: 8
US-10-030-194A-4 (1-572) x AAD06646 (1-1951)
Qy 5 LeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySerThrSerSer 24
Db 171 ATCCATCTCTGAAAAAACCACCATGAGAGAGATCATCATCATCATCATCA----- 224
Qy 25 ProAlaValPheGlyLys-AspLysMetMetValLysGluGluGluAAsp----- 41
Db 225 -----AGATAGAAGACTATGATGATGAATGAAGAAGACGCGTAACGG 269
Qy 42 ----AspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVa 60
Db 270 CATGATGAGGCTTCTAGCTGTTCTTGGTTACAGGTAGTTCGCGAAATGGCTGATGT 329
Qy 60 lAlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAl 80
```

Db 330 TCTCAGAACTCGAGACCTTTGAAGTTATGATGCTTAATGTTCAAGAGACGATCTTTC 389
Qy 80 aHisLeuAlaThrAspThrValHisTyAsnProAlaGluLeuTySerTrpLeuAspAs 100
Db 390 TCAACTCGCTACTGAGACTGTTCACTATAATCCGGGAGCTTTACACGTGGCTTGATTC 449
Qy 100 mMetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGl 120
Db 450 TATGCTCCAGCACCTTAATCTCTCGCTCT-----AACCGCA 488
Qy 120 uIleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLeuAlaIleProGl 140
Db 489 GTAC-----GATCTTAAGCTATTCCCGG 512
Qy 140 yAsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSer----- 158
Db 513 TGACGGATTCTC-----AATCAGTTCCGTATCGATTCGGCTCTCTCGCTAACCA 563
Qy 159 -----AsnLysArgLeuLysProSerSerSerProAs 169
Db 564 AGCGCGCGAGAGATACGTATACAAACAAAGCGTTGAAATGCTCAAAAC----- 615
Qy 169 pSerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValTh 189
Db 616 -----GGCGTCGTGGAACACCC-----ACAGCGACCGC 644
Qy 189 rGluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyValArgLeuValHi 209
Db 645 TGAGTCAACTCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704
Qy 209 sAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLe 229
Db 705 CGCGCTTTGGCTTCGGCTGAAGCTGTTTCAAGAGAGAACTGCTGCTGGCGGAAGCTCT 764
Qy 229 uValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaTh 249
Db 765 GGTGAAGCAAAATCGGATCTTGTAGCTGTTTCTCAAAATCGGAGCTATGAGCAAGTCGCTAC 824
Qy 249 rTyrrPheAlaGluAlaLeuAlaAlaArgArgIleTyrrArgLeuSerProGlnThrGlnI 269
Db 825 TTACTTCGCGGAGCTTCGCGCGCGGATTTACCGTCTCTCTCGCTCGCAGAGTCCAA 884
Qy 269 eAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrrGluThrCysProTyrrLeuLy 289
Db 885 CGACCACTCTCTCCGATACCTCTTCAGATGACCTTCACGAGACTTGTCTCTTATCTCAA 944
Qy 289 sPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgVa 309
Db 945 GTTCGCTCACTTCACGGCGAATCAAGCGATTCTCGAAGCTTTTCAAGGGGAGAAAGAGT 1004
Qy 309 lHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLe 329
Db 1005 TCAUGTCATTGATTTCTTATGAGTCAAGGTTTCAATGGCGCGCGCTTATGCGAGCTCT 1064
Qy 329 uAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAl 349
Db 1065 TGGCTTCGACCTGGTGGTCTCTCTGTTTCCGGTTAACCGGAATTTGTCACCGGCACC 1124
Qy 349 aAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaI 369
Db 1125 GGATAAATTCGATTAATCTTATGAGTGAAGTGGGTGTAAGCTGGCTCATTTAGCTGAGCGCAT 1184
Qy 369 eHisValGluPheGluTyrrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSe 389
Db 1185 TCAGTTGAGTTGAGTACAGAGGATTTGTGCTAACCACTTGTAGCTGATCTTGTAGCTTC 1244
Qy 389 rMetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLe 409
Db 1245 GATGCTTGAGCTTAGACCAAGTGAATGTAATCTGCTGGGTAACTCTGTTTCGAGCT 1304
Qy 409 uHisLysLeuLeuGlyArgThrGlyGlyIleGluLysValPheGlyValValLysGlnI 429

Db 1305 TCACAAAGCTCTTGGGACACCTGGTGGCATCGATGAAGGTTCTTGTGTGGTGAATCAGAT 1364
Qy 429 eLysProValIlePheThrValValGluGlnGlnSerAsnHisAsnGlyProValPheLe 449
Db 1365 TAAACCGGAGATTTTCACTGTGGTTGAGCAGGAATCGAACCAATAATAGTCCGATTTCTT 1424
Qy 449 uAspArgPheThrGluSerLeuHisTyrrTyrrSerThrLeuPheAspSerLeuGluGlyAl 469
Db 1425 AGATCGGTTTACTGAGTCGTTGCATATTACTCGACGTTGTTTGTACTCGTTGAAGGTGT 1484
Qy 469 aProSerSerGlnAspLysValMetSerGluValTyrrLeuGlyLysGlnIleCysAsnLe 489
Db 1485 ACCGAGTGGTCAGACAAAGGTCATGTCGAGGTTTACTTGGTAAACAGATCTGCAAGT 1544
Qy 489 uValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAs 509
Db 1545 TGTGGCTTGTGATGACCTGACCGAGTTGAGCGTTCATGAACGTTGAGTCAGTGGAGGAA 1604
Qy 509 nArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAl 529
Db 1605 CCGGTTCCGGGCTCTGCTGGGTTTGGCGGCTGSCACATATGCTTGAATGCGTTTAAACAAGC 1664
Qy 529 aSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrrArgValGluLysAsnAsnGl 549
Db 1665 GAGTATGCTTTTGGCTCTGTTCAACGCGGCTGAGGTTATCGGGTGGAGAGGTGACGG 1724
Qy 549 yCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAlaTrpLysLeuSe 569
Db 1725 CTGCTCATGTTGGTTGGCACACACGACCGCTCATACCCACCTCGGCTTGGAACTCTC 1784
Qy 569 r 569
Db 1785 C 1785
RESULT 12
ADD55687
ID ADD55687 standard; cDNA; 1951 BP.
XX
AC ADD55687;
XX
DT 15-JAN-2004 (first entry)
XX
DE Thalecress environmental stress-related cDNA #23.
XX
KW Thalecress; environmental stress; ss; gene; plant; viral infection;
KW fungal infection; microbial infection; herbicide resistance; heat; cold;
KW heavy metal; low light; drought; osmotic stress; salt concentration;
KW transgenic.
XX
OS Arabidopsis thaliana.
XX
PN US2003131386-A1.
XX
PD 10-JUL-2003.
XX
PF 22-OCT-2002; 2002US-00278536.
XX
PR 23-MAR-1999; 99US-0125814P.
PR 22-MAR-2000; 2000US-00532591.
XX
PA (SAMA/) SAMAH R.
PA (HEAR/) HEARD J.
PA (JIAN/) JIANG C.
PA (PINE/) PINEDA O.
PA (REUB/) REUBER L.
PA (RIEC/) RIECHMANN J L.
PA (YUGG/) YU G.
PA (KEDD/) KEDDIE J.
PA (RATC/) RATCLIFFE O.
PA (PILG/) PILGRIM M.
PA (ADAM/) ADAM L.
PA (BROU/) BROUN P.
PA (ZHAN/) ZHANG J.

XX Samaha R, Heard J, Jiang C, Pineda O, Reuber L, Riechmann JL;
PI Yu G, Keddie J, Ratcliffe O, Pilgrim M, Adam L, Broun P, Zhang J;
XX WPI; 2003-829608/77.
DR P-PSDB; ADD55688.
XX
PT New recombinant polynucleotide for altering the regulation of gene
PT expression of plants to modify the plant's traits, particularly the
PT plant's environmental stress tolerance.
XX
PS Claim 14; SEQ ID NO 45; 219pp; English.
XX
CC The invention relates to a recombinant polynucleotide that alters a
CC plant's environmental stress tolerance when compared with the same trait
CC of another plant lacking the recombinant polynucleotide. Also included
CC are a transgenic plant comprising the novel recombinant polynucleotide
CC having a sequence that encodes a polypeptide comprising at least 6
CC consecutive amino acids of any of the 55 250-500 residue amino acid
CC sequences (S1), given in the specification, altering the environmental
CC stress response or tolerance of a plant, or altering a plant's trait and
CC altering the expression levels of at least one gene in a plant. The
CC recombinant polynucleotide and methods are useful for altering the
CC regulation of gene expression of plants to modify the plant's traits, in
CC particular with respect to environmental stress responses (e.g. to viral
CC infection, fungal infection, microbial infection, herbicide resistance,
CC heat, cold, heavy metals, low light, drought, osmotic stress and salt
CC concentration). The present sequence is an environmental stress related
CC polynucleotide of the invention.
XX
SQ Sequence 1951 BP; 482 A; 419 C; 475 G; 575 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.37e-180 Length: 1951
Score: 2106.50 Matches: 428
Percent Similarity: 80.55% Conservative: 40
Best Local Similarity: 73.67% Mismatches: 55
Query Match: 71.41% Indels: 58
DB: 10 Gaps: 8

US-10-030-194A-4 (1-572) x ADD55687 (1-1951)

QY 5 LeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySerThrSerSer 24
DB 171 ATCCATCTCTGAAAACCCACCATGACGAGAGATCATCATCATCATCA----- 224

QY 25 ProAlaValPheGlyLys-AspLysMetMetMetValLysGluGluGluAsp----- 41
DB 225 -----AGATAAGAAGACTATGATGATGAATGAAGAAGACGCGTAAACGG 269

QY 42 ----AspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVa 60
DB 270 CATGATGAGCTTCAGCTGTTCTGGTTACAGGTTAGGTCATCGGAATAGGCTGATGT 329

QY 60 lAlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAl 80
DB 330 TGCTCAGAACTCGAGAGCTTGAGTTATGATGCTAATGTTCAAGAAGACGATCTTC 389

QY 80 aHisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTrpLeuAspAs 100
DB 390 TCAACTCGCTACTGAGACTGTTCATATAATCCGCGGAGCTTTACACGTCGCTGATTC 449

QY 100 nMetLeuThrGluLeuAsnProAlaAlaThrThrGlySerAsnAlaLeuAsnProGl 120
DB 450 TATGCTCACCGACCTTAATCTCCGCTCT-----AAGCCGA 488

QY 120 uIleAsnAsnAsnAsnAsnAsnSerPhePheThrGlyGlyPheLeuLysAlaIleProGl 140
DB 489 GTAC-----GATCTTAAGCTATTCCCG 512

QY 140 yAsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSer----- 158
DB 513 TGACGCGATTCTC-----AATCAGTTGCTATCATCGGCTTCTTCGTTCAACCA 563

QY 159 -----AsnLysArgLeuLysProSerSerSerProAs 169
DB 564 AGCGCGGAGGAGATAGTATATACAAACACGCGTTGAATGCTCAAC----- 615

QY 169 pSerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValTh 189
DB 616 -----GGCGTCTGGMAAACACCAC-----ACAGCGACGCGC 644

QY 189 rGluSerThrArgProLeuLeuValAspSerGlnAspAsnGlyValArgLeuValHi 209
DB 645 TGAGTCAACTCGCGCATGTTGCTGGTGGTACTCGCAGGGAACGGTGTGCTCTCGTTCA 704

QY 209 sAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLe 229
DB 705 CGCGCTTTGGCTTCGCTGAGAGCTGTTCAAGAGAGANCTGACTGTGGCGGAAGCTCT 764

QY 229 uValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaTh 249
DB 765 GGTGAAGCAAAATCGGATTTCTAGCTGTTCTCAATCGGAGCTATGAGACAAAGTCGCTAC 824

QY 249 rTyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProGlnThrGlnIl 269
DB 825 TTACTTCCCGAAGCTCTCGCGCGCGGATTTACCGTCTCTCTCGTCGAGAGTCCAAAT 884

QY 269 eAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLy 289
DB 885 CGACCATCTCTCTCCGATACTCTTCAGATGCACCTTCACGAGACTTGTCTTATCTCAA 944

QY 289 sPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgVa 309
DB 945 GTTCGCTCACTTCACGCGGAATCAAGCATTTCTCGAAGCTTTTCAAGGGAAGAAAGAGT 1004

QY 309 lHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLe 329
DB 1005 TCATGTCTATGATTTCTCTATGAGTCAAGGCTTTCATATGGCGCGGCTTATGCGAGCTCT 1064

QY 329 uAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAl 349
DB 1065 TGCCTTCGACCTGGTGGTCTCTCTGTTTCCGGTTAACCGGAATTTGGTCCACCGGACC 1124

QY 349 aAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIl 369
DB 1125 GGATAATTTTCGATTTATCTTCATGAAGTTGGTGTGAAGCTGGCTCATTTAGCTGAGGCGAT 1184

QY 369 eHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSe 389
DB 1185 TCACGTTGAGTTTGAGTACAGAGGATTTGTGGCTAACACACTTTAGCTGATCTTGATGCTTC 1244

QY 389 rMetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLe 409
DB 1245 GATGCTTGAGCTTAGACCAAGTGAATGATCTGTTGGGTTAACTCTGTTTTCGAGCT 1304

QY 409 uHisLysLeuLeuGlyArgThrGlyIleGluLysValPheGlyValValLysGlnIl 429
DB 1305 TCACAAGCTCTTGGGACGACCTGGTGGATCGATCGATAAGGTTCTTGGTGTGGTGAATCAGAT 1364

QY 429 eLysProValIlePheThrValValGlnGluSerAsnHisAsnGlyProValPheLe 449
DB 1365 TAAACCGGAGATTTTCATCTGGTGGTGGACGAGAAATCGAACCATATAGTCCGATTTCTT 1424

QY 449 uAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAl 469
DB 1425 AGATCGGTTTACTGAGTCTGTTGCATTTACTTCGACGTTGTTGTCATCGTTGGAAGGTGT 1484

QY 469 aProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLe 489
DB 1485 ACCGAGTGGTCAAGACAAAGGTCATGTCGAGGTTTACTTGGGTTAAACAGATCTGCAACGT 1544

QY 489 uValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAs 509
DB 1545 TGTGGCTTGTGATGGACCTGACCGAGTTGAGCGTCATGAACGTTGAGTCAGTGAGGAA 1604

QY 509 nArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAenAlaPheLysGlnAl 529
Db 1605 CCGGTTCCGGCTCTGGGCTTTGGCGCTGCACATATGGTTGCGAATGCGTTTAAGCAAGC 1664
QY 529 aSerThrLeuLeuAlaLeuPheAenGlyGlyGluGlyTyrArgValGluLysAenAenG1 549
Db 1665 GAGTATGCTTTTGGCTCTGTTCAACGGCGGTGAGGTTATCGGTGGAGGAGTGACGG 1724
QY 549 yCysLeuMetLeuSerTrpHisThrArgProLeuileThrThrSerAlaTrpLysLeuSe 569
Db 1725 CTGCTCATGTTGGGTTGGCACACAGCGCTCATAGCCACCTCGGCTTGGAAACTCTC 1784
QY 569 r 569
Db 1785 c 1785
RESULT 13
ADD30751
ID ADD30751 standard; cDNA; 1951 BP.
XX AC
XX ADD30751;
XX
XX 15-JAN-2004 (first entry)
XX Plant yield-related polynucleotide clone G308.
XX
XX ds; transcription factor; transgenic plant; growth rate; senescence;
XX seed germination rate; plant vigor; seedling vigor.
XX Arabidopsis thaliana.
XX
XX WO2003013227-A2.
XX
XX PD 20-FEB-2003.
XX
XX 09-AUG-2002; 2002WO-US025805.
XX
XX 09-AUG-2001; 2001US-0310847P.
XX 19-NOV-2001; 2001US-0336049P.
XX 11-DEC-2001; 2001US-0338692P.
XX 14-JUN-2002; 2002US-00171468.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
XX Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
XX Broun PE;
XX
XX WPI; 2003-248221/24.
XX P-PSDB; ADD30752.
XX
XX New plant transcription factor polynucleotides and polypeptides, useful
XX in producing transgenic plants with commercially valuable properties,
XX such as an alteration in a plant growth characteristic, e.g. growth rate
XX or apomixis.
XX
XX Disclosure; SEQ ID NO 780; 454pp; English.
XX
XX The invention relates to a number of isolated Arabidopsis thaliana cDNA
XX sequences and their encoded proteins which are especially transcription
XX factor related cDNA's and proteins. The isolated or recombinant plant
XX transcription factor polynucleotides and polypeptides are useful in
XX producing transgenic plants with commercially valuable properties, i.e.
XX modified or altered desirable traits as compared to a reference plant,
XX such as an alteration in a plant growth characteristic, e.g. growth rate,
XX germination rate of seeds, vigor of plants and seedlings, or leaf and
XX flower senescence. Sequence information related to the polynucleotides
XX and polypeptides can also be used in bioinformatic search methods. The
XX transgenic plant is useful for growing a progeny plant from a parent
XX plant. This sequence represents one of the cDNAs of the invention.
XX
XX Sequence 1951 BP; 482 A; 419 C; 475 G; 575 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 1.37e-180 Length: 1951
Score: 2106.50 Matches: 428
Percent Similarity: 80.55% Conservative: 40
Best Local Similarity: 73.67% Mismatches: 55
Query Match: 71.41% Indels: 58
DB: 10 Gaps: 8
US-10-030-194a-4 (1-572) x ADD30751 (1-1951)
QY 5 LeuHisGlnPheGlnGlyProAenHisGlyThrSerIleAlaGlySerSerThrSerSer 24
Db 171 ATCCATCTCTCAAAAACCCACCATGAGAGAGATCATCATCATCATCA----- 224
QY 25 ProAlaValPheGlyLys-AspLysMetMetMetValLysGluGluGluAsp----- 41
Db 225 -----AGATAAGAAGACTATGATGATGAATGAAGAAGACGCGTAACGG 269
QY 42 ----AspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVa 60
Db 270 CATGGATGAGCTTCTAGCTGTTCTTGGTTACAAGGTTAGGTCATCGAAATGGCTGATGT 329
QY 60 lAlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAenAlaGlnGluAspGlyLeuAl 80
Db 330 TGCTCAGAACTCGAGCAGCTTGAAGTTATGATGCTTAATGTTCAAGAAAGACGATCTTTC 389
QY 80 aHisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTrpLeuAspAs 100
Db 390 TCAACTCGCTACTGAGACTGTTCACTATAATCCGGCGGAGCTTTACACGCTGGCTGATTC 449
QY 100 nMetLeuThrGluLeuAenProProAlaAlaThrThrGlySerAenAlaLeuAenProGl 120
Db 450 TATGCTCACCAGCCTTAATCTCTCGTCGCT-----AACGCCGA 488
QY 120 uIleAenAsnAenAenAenSerPhePheThrGlyAspLeuLysAlaIleProGl 140
Db 489 GTAC-----GATCTTAAAGCTATTCCCG 512
QY 140 yAenAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerSer----- 158
Db 513 TGACGCGATTCTC-----AATCAGTTCTGCTATCGGCTTCTTCGCTAACCA 563
QY 159 -----AsnLysArgLeuLysProSerSerSerProAs 169
Db 564 AGCGCGCGGAGAGATAGTATACTACAAACAAGCGTTGAATGCTCAAC----- 615
QY 169 pSerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValTh 189
Db 616 -----GGCGTCTGGAAACCACC-----ACGCCGCCG 644
QY 189 rGluSerThrArgProLeuLeuLeuValAspSerGlnAspAenGlyValArgLeuValHi 209
Db 645 TGAGTCAACTCGCATGTTGCTCTGTTGACTCCGAGGAGAACCGTGTGCGTCTCGTTCA 704
QY 209 sAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAenLeuThrLeuAlaGluAlaLe 229
Db 705 CGCGCTTTTGGCTTCGCTGAAGCTGTTCAAGAAGAGAAATCTGACTGTGGCGGAAGCTCT 764
QY 229 uValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaTh 249
Db 765 GGTGAAGCAATCGGATCTTAGCTGTTTCTCAATCGAGAGCTATGAGACAAAGTCGCTAC 824
QY 249 rTyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProProGlnThrGlnI 269
Db 825 TTACTTCGCCGAAGCTCTCGCGCGGCGGATTTACCGTCTCTCTCGTCGAGAGTCCAAT 884
QY 269 aAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuL 289
Db 885 CGACCACTCTCTCCGATACCTCTTACATGACCTTCTACGAGACTTCTCTCTTATCTCAA 944
QY 289 sPheAlaHisPheThrAlaAenGlnAlaIleLeuGluAlaPheGluGlyLysLysArgVa 309
Db 945 GTTCGCTCACTTCACGGCGAATCAAGCGATTCTCGAAGCTTTTCAAGGGAAGAAAGAGT 1004

Qy 309 lHisValileAaspPheSerMetAsnGlnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLe 329
|||||
Db 1005 TCATGTCATTGATTTCTTATGAGTCAGAGCTTTCATATGGCGGCTTATGCAGGCTCT 1064

Qy 329 uAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAl 349
|||||
Db 1065 TGCCTTCGACCTGGTGGTCTCTCTGTTTTCGGTTAAACGGAAATTTGGTCCACCGGCACC 1124

Qy 349 aAspAsnSerAspHisLeuHisGluValGlyCysIysLeuAlaGlnLeuAlaGluAlaIl 369
|||||
Db 1125 GGATAAATTTCCGATATCTTCATGAAGTTGGGTGAAGCTGGCTCATTTAGCTGAGGCGAT 1184

Qy 369 eHisValGluPheGluTyraGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSe 389
|||||
Db 1185 TCACGTTGAGTTGAGTACAGAGGATTTGTGGCTTAACACTTGTGATCTTGATGCTTC 1244

Qy 389 rMetLeuGluLeuArgProSerGluThrGluAlaValAlaValaAsnSerValPheGluLe 409
|||||
Db 1245 GATGCTTCGAGCTTAGACCAAGTAGAGATTGAATCTGTTGCGGTTAACTCTGTTTCGAGCT 1304

Qy 409 uHisIysLeuLeuGlyArgThrGlyGlyIleGluIysValPheGlyValIysGlnIl 429
|||||
Db 1305 TCACAAGCTCTTGGGACACCTGGTGGCATGATAAGGTTCTTGGTGTGGTGAATCAGAT 1364

Qy 429 eLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPheLe 449
|||||
Db 1365 TAAACCGAGATTTTCATCTGGTTGACAGAGNATCGAACCATATAGTCCGATTTCTT 1424

Qy 449 uAspArgPheThrGluSerLeuHisTyrrSerThrLeuPheAspSerLeuGluGlyAl 469
|||||
Db 1425 AGATCGGTTTACTAGTCGTTGCATTATTAATCTCGACGTTGTTGACTCGTTGGAAGGTGT 1484

Qy 469 aProSerSerGlnAspLysValMetSerGluValTyrrLeuGlyIysGlnIleCysAsnLe 489
|||||
Db 1485 ACCGAGTCGTCAAGACAAGGTCATGTCGGAGGTTTACTTGGGTAAACAGATCTGCAACGT 1544

Qy 489 uValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAs 509
|||||
Db 1545 TGTGGCTTGTATGACACCTACCGAGTTGAGCGTCATGAACGTTGAGTCAGTGAGGAA 1604

Qy 509 nArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAl 529
|||||
Db 1605 CCGGTTTCGGTCTGCTGGGTTTGGCGCTGCACATATTGTTTCCGATGCGTTTAAAGCAAGC 1664

Qy 529 aSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrrArgValGluLysAsnAsnGl 549
|||||
Db 1665 GAGTATGCTTTTGGCTCTGTTCAACGCGGCGTGGAGGTTTATCGGTTGGAGAGAGTGACGG 1724

Qy 549 yCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAlaTrpLysLeuSe 569
|||||
Db 1725 CTGTCATGTTGGTTGGTCACACAGACCGCTCATAGCCACCTCGGCTTGGAAACTCTC 1784

Qy 569 r 569
Db 1785 C 1785

RESULT 14
ADI43826
ID ADI43826 standard; DNA; 1951 BP.
XX
AC ADI43826;
XX
DT 22-APR-2004 (first entry)
XX
DE Plant transcription factor related polynucleotide #1502.
XX
KW transgenic; plant; enhanced tolerance to abiotic stress;
KW glyphosphate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;
KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;

KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
transcription factor; ds.
XX
OS Unidentified.

XX US2004019927-A1.
XX 29-JAN-2004.

XX 25-FEB-2003; 2003US-00374780.
XX 18-APR-2001; 2001US-00837944.

XX (SHER/) SHERMAN B K.
PA (RIEC/) RIECHMANN J L.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J E.
PA (HAAK/) HAAKE V.
PA (CREE/) CREELMAN R A.
PA (RATC/) RATCLIFFE O.
PA (ADAM/) ADAM L J.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIS J.
PA (BROU/) BROUN P E.
PA (PILG/) PILGRIM M L.
PA (DUBE/) DUBELL A N.
PA (PINE/) PINEA O.
PA (YUGG/) YU G.

XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX WPI; 2004-132245/13.
DR P-PSDB; ADI43827.

XX New transgenic plant comprising a recombinant polynucleotide of any one
of more than 500 nucleotide sequences, useful in bioinformatic search
methods.

XX Disclosure; SEQ ID NO 2289; 435pp; English.

The invention describes a transgenic plant comprising a recombinant
polynucleotide of any one of more than 500 nucleotide sequences fully
defined in the specification or its complement. The method of the
invention can be used to produced a plant having altered traits such as:
enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
sensitivity; disease resistance; sugar sensing; early or late flowering;
altered flower structure, change in stem bifurcations, altered branching
pattern, reduced apical dominance, reduced trichome density; lack of
trichomes; reduced ectopic trichome development; altered trichome
development; increase in trichome number; altered stem morphology;
increased root growth; increased root hair; altered seed development;
altered cell proliferation or cell differentiation; rapid development;
premature senescence; increased necrosis; increase in seedling or plant
size; decreased plant size; leaf morphology; seed morphology; seed
biochemistry; increase in root anthocyanins; increase in plant
anthocyanins, or alteration in light response or shade avoidance. The
transgenic plant, polynucleotides and polypeptides are useful in
bioinformatic search methods. This sequence represents a plant
transcription factor related polynucleotide.

XX Sequence 1951 BP; 482 A; 419 C; 475 G; 575 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.37e-180 Length: 1951
Score: 2106.50 Matches: 428
Percent Similarity: 80.55% Conservative: 40
Best Local Similarity: 73.67% Mismatches: 55
Query Match: 71.41% Indels: 58
DB: 12 Gaps: 8

Qy	5	LeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySerSerThrSerSer	24
Db	171	ATCCATCTCTGAAACCAACCAACCATGAAGAGAGATCATCATCATCATCATCA	224
Qy	25	ProAlaValPheGlyLys-AspLysMetMetMetValLysGluGluGluAsp	41
Db	225	-----AGATAAGAAGACTATGATGATGAATGAAGAAGACGACGGTAACGG	269
Qy	42	---AspGluLeuLeuGlyValLeuGlyTyrLysValArgSerSerGluMetAlaGluVa	60
Db	270	CATGGATGAGCTTCTAGCTGTCTTGTTTACAAAGTTAGTGCATCGAANAATGGCTGATGT	329
Qy	60	lAlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAl	80
Db	330	TGCTCAGAAACTCGAGCAGCTTGAAGTTATGATGCTAATGTTCAAGAAGACGATCTTTC	389
Qy	80	aHisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTyrLeuAspAs	100
Db	390	TCAAATCGCTACTAGACTGTTCACATAAATCCGGCGGAGCTTTACACGTGCTTGATTC	449
Qy	100	nMetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGl	120
Db	450	TATGCTCACCGACCTTAATCTCCGTCGTCT-----AACGCCGA	488
Qy	120	uIleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGl	140
Db	489	GTAC-----GATCTTAAAGCTATTCCTCCGG	512
Qy	140	yAsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerSer	158
Db	513	TGACGCGATTCTC-----AATCAGTTCTGATTCGATTCGGCTTCTTCGCTTAACCA	563
Qy	159	-----AsnLysArgLeuLysProSerSerSerSerSerProAs	169
Db	564	AGCGCGCGGAGAGATACGTATATCAACAACAAAGCGTTGAAATGCTCAAC	615
Qy	169	pSerMetValThrSerProSerProAlaGlyValIleGlyThrThrValThrValThr	189
Db	616	-----GGCGTCGTGGNAACCAAC-----ACAGCAGCGC	644
Qy	189	rGluSerThrArgProLeuIleLeuValAspSerGlnAspGlyValArgLeuValHi	209
Db	645	TGAGTCAACTCGCGATGTGCTGCTGCTGACTCGCAGGAGAACGGTGTGCGTCTCGTTCA	704
Qy	209	sAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLe	229
Db	705	CGCGCTTTTGGCTTCGCGTGAAGCTGTTTCAGAAAGAGAAATCTGACTGTGGCGGAAGCTCT	764
Qy	229	uValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaTh	249
Db	765	GGTGAAAGCAATCGGATTTCTAGCTGTTTCTCAAATCGGAGCTATGAGACAAGTGCCTAC	824
Qy	249	rTyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProProGlnThrGlnI	269
Db	825	TTACTTCGCCGAAGCTCTCGCGCGCGGATTTACCGTCTCTCTCCGTCGCAGAGTCCAAT	884
Qy	269	eAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLy	289
Db	885	CGACCACCTCTCTCCGATACTCTTCAGATGCACCTTCACGAGACTGTGCTTATCTCA	944
Qy	289	sPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGlyLysArgVa	309
Db	945	GTTTCGCTCACTTCAGCGGAATCAAGCGAATCTCGAAGCTTTTCAAGGGAAGAAAGAGT	100
Qy	309	lHisValIleAspPheSerMetAsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLe	329
Db	1005	TCATGTCATTATTCTATGAGTCAAGTCTTCAATGGCCGGCGCTTATCGAGCTCT	1061
Qy	329	uAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGlyProProAlaAl	349
Db	1065	TGCGCTTCGACCTGGTGGTCTCTCTGTTTTCGGTTAACCGAATGGTTCACCGGACCC	1121

Qy	349	aAspAmsSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaI	369
Db	1125	GGATTAATTTTCGATTATCTTCATGAAGTTGGGTGAAGCTGGCTCATTTAGCTGAGGCGAT	1184
Qy	369	eHisValGluPheGluTyrrArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSe	389
Db	1185	TCACGTTGAGTTTGAGTACAGAGATTTTGCGTAAACCTTTAGCTGATCTTGATGCTTC	1244
Qy	389	rMetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSerValPheGluLe	409
Db	1245	GATGCTTTGAGCTTAGACCAAGTCAGATTGAATCTGTTGCGTTAACTCTGTTTCGAGCT	1304
Qy	409	uHisLysLeuLeuGlyArgThrGlyGlyLleGluLysValPheGlyValValLysGlnI	429
Db	1305	TCACACAGCTCTTGGGACGACCTCGTGGCATCCGATTAAGGTTCTTGTTGTGTGGAATCAGAT	1364
Qy	429	eLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGlyProValPheLe	449
Db	1365	TAAACCGGAGATTTTCACTGTGTTGACGAGGAATCGAACCAATAATAGTCGATTTCTT	1424
Qy	449	uAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAl	469
Db	1425	AGATCGGTTTACTGAGTCGTTGCATTATTACTCGACGTTCTTTGACTCGTTGGAAGGTGT	1484
Qy	469	aProSerSerClnAspLysValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLe	489
Db	1485	ACCAGGTGGTCACAGACAAGGTCACTGCGGAGGTTTACTTGGGTAAACAGATCTGCAACGT	1544
Qy	489	uValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSerGlnTrpSerAs	509
Db	1545	TGTGGCTTTGTGATGGACCTCACCAGTTGAGCGTCATGAACGTTGAGTCAGTGGAGGAA	1604
Qy	509	nArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAl	529
Db	1605	CCGGTTCCGGGTCTGCTCGGTTTCGGCTGCACATATTGTTTCGAATTCGTTTAAAGCAAGC	1664
Qy	529	aSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGl	549
Db	1665	GAGTATGCTTTTGCTCTGTTTCAACGCGGTGAGGGTTATCGGGTGAGGAGTAGTCAGG	1724
Qy	549	yCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAlaTrpLysLeuSe	569
Db	1725	CTGTCATGTTGGGTTGGCACACACGACCGCTCATAGCCACCTCGGCTTTGGAACTCTC	1784
Qy	569	r 569	
Db	1785	C 1785	
RESULT 15			
AD001804			
ID	AD001804	standard; cDNA; 1951 BP.	
AC	AD001804;		
XX			
DT	01-JUL-2004	(first entry)	
DE	Thalassia transcription factor cDNA #109.		
XX			
KW	Thalassia; transcription factor; ss; gene; plant; transgenic;		
KW	abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;		
KW	phosphate limitation; potassium limitation; nitrogen limitation;		
KW	hormone sensitivity; disease resistance; sugar sensing; seed germination;		
KW	flowering; inflorescence architectural change;		
KW	meristem cell differentiation; phyllotaxy; apical dominance;		
KW	trichome development; seed development; premature senescence;		
KW	delayed senescence; lethality; necrosis; plant size; leaf morphology;		
KW	seed morphology; secondary metabolism; light response; shade avoidance.		
OS	Arabidopsis thaliana.		
XX			
FN	US2004045049-A1.		
XX			

THIS PAGE BLANK (USP10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 2, 2004, 19:11:59 ; Search time 9414 Seconds
(without alignments)
2873.349 Million cell updates/sec

Title: US-10-030-194A-4
Perfect score: 2950
Sequence: 1 MKKDLHQFQPNHGTSTAGS.....LSMHTPLITTSAAKLSAVH 572

Scoring table:

BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10030194/runat_01112004_184752_28637/app.query.fasta_1.711
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -WATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOFCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRP=US10030194 @CGN 1 1 7406 @runat_01112004_184752_28637 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.btg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	2950	100.0	1779 6 AX081278
2	2946	99.9	1779 6 AX081276
3	2408.5	81.6	2210 8 ATRGA1
4	2407	81.6	1764 6 AX507818

ALIGNMENTS

RESULT 1	AX081278	Sequence 3 from Patent WO0109356.	AX081278	1779 bp	DNA	linear	PAT 27-FEB-2001
LOCUS	AX081278	Sequence 3 from Patent WO0109356.	AX081278	1779 bp	DNA	linear	PAT 27-FEB-2001
DEFINITION	AX081278	Sequence 3 from Patent WO0109356.	AX081278	1779 bp	DNA	linear	PAT 27-FEB-2001
ACCESSION	AX081278	Sequence 3 from Patent WO0109356.	AX081278	1779 bp	DNA	linear	PAT 27-FEB-2001
VERSION	AX081278.1	GI:13170127	AX081278.1	GI:13170127			
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
FEATURES							
source							
CDS							

BT010467 Arabidops
AY054160 Arabidops
AC005560 Arabidops
AY052239 Arabidops
Y15194 Arabidops
AY150331 Sequence
AY142002 Arabidops
AY058194 Arabidops
AC006917 Genomic s
A64697 Sequence 1
AR174879 Sequence
AR252361 Sequence
Y15193 Arabidops
Y11336 A.thaliana
A64699
A64699 Sequence 3
AR174880 Sequence
AR252362 Sequence
A64705 Sequence 9
AR174883 Sequence
AR252365 Sequence
A64701 Sequence 5
AR174881 Sequence
AR252363 Sequence
AF378125 Vitis vin
A64703 Sequence 7
AR174882 Sequence
AR252364 Sequence
AY326307 Cucurbita
AY326306 Cucurbita
AY269087 Lycopersi
AY208992 Gossypium
AJ242530 Zea mays
AX005806 Sequence
BD074480 Genetic r
AB030956 Oryza sat
AC087797 Oryza sat
AF460219 Hordeum v
AJ242531 Triticum
AX005805 Sequence
BD074479 Genetic r
AF492572 Argemone

Renard, M., Delourme, R., Barret, P., Brunel, D., Froger, N. and Tanguy, X.
Mutant gene of the gras family and plants with reduced development containing said mutant gene
Patent: WO 0109356-A 3 08-FEB-2001;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR)
Location/Qualifiers
1..1779
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"
60..1778
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC33121.1"

/db xref="GI:13170128"
/translations="MKRDLHQFOGPNHGTSTIAGSSSTSPAVFGKDKMMVKEREDDEL
LGVLYKVRSEAEVALKLEQLFTVMGNAQEDGLAHATDTVHYNPAELYSLDNML
TELPAAATIGSNALNPEINNNNSFTGDLKALPGNAVCRRSNQFAFAYDSSNK
RLKPSSPDSMTSPSPAGVGTGTTVTVTETSTRLILVDSQDNGVRLVHALMACAAV
QSNLTLAEALVQIGFLAVSQAGAKVKATYFAELARRIYRLSPQTOIDHSLADT
LOMHFYETCYLFAFHAFNDQALILEAPEGKKRVHVIDFSMNOGLQWALMALAREG
GPSPRLTGTGPPAADNDHLEHVGCKLAOLAEAIHVEPRYGFVANSIADLDSMLE
LRPSEYAVAVNSVPELHKLKLGRTGGIEKVGUVKQIKPVIPTVVEQESHNGPVLID
RFTESLHYSTLDSLEGAPSSQKVNSEVTLGKQICNLVACSGPDRKVERHETLSQWS
NRFSGSGFAHGLGSNAFKAOSTLLALFNGEGGYRVERKNGCLMLSWHTRPLITTSAM
KLSAVH"

ORIGIN

Alignment Scores:
Pred. No.: 7,92e-218 Length: 1779
Score: 2950.00 Matches: 572
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-4 (1-572) x AX081276 (1-1779)

Qy	1	MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer	20
Db	60	ATGAGAGGATCTTCAATCAGTTCAGGTCCCAACCCAGGACATCAATCGCGGTTCT	119
Qy	21	SerThrSerProAlaValPheGlyLeuAspLysMetMetValLysGluGlu	40
Db	120	TCCACTCTTCCCCTGCGGTGTTTGGTAAAGCAAGATGATGATGGTCAAAAGAAAGAA	179
Qy	41	AspAspGluLeuLeuGlyValLeuGlyTyrIlyValArgSerSerGluMetAlaGluVal	60
Db	180	GAGCAGAGCTTCTAGAGGCTTGGGTTACAAGGTTAGGCTTCGGAGATGGCTGAGGTT	239
Qy	61	AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla	80
Db	240	CGGTGGAACCTCGAGCAGCTTGAGACGATGATGGTTAACGCTCAAGAGACGGTTAGCT	299
Qy	81	HisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSerTrpLeuAspAsn	100
Db	300	CACCTCGCAGCGATACGTGTTCAATCAACCCCGCTGAGCTTTACTCGTGGCTTGATAAC	359
Qy	101	MetLeuThrGluLeuAsnProProAlaAlaThrGlySerAsnAlaLeuAsnProGlu	120
Db	360	ATGCTCAGGAGCTTAACCCACCGCTGCAACCGCGATCTAACGCTTTGAAACCCCGGAG	419
Qy	121	IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly	140
Db	420	ATTATATATATATATATACTCGTTTTTCAACCGAGCGACCTCAAGCGATTCCTGGA	479
Qy	141	AsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSerSerAsnLys	160
Db	480	AACCGCGTTGTCGACATCTAATCAGTTCGCGTTTCGCGTTGATTCTCGAGTAATAAG	539
Qy	161	ArgLeuLysProSerSerSerProAspSerMetValThrSerProSerProAlaGlyVal	180
Db	540	CGTTTGAAACCGCTCTCGAGCCCTGATTTCATGGTTACATCTCCATCACCCTGCGGAGTT	599
Qy	181	IleGlyThrThrValThrValThrGluSerThrArgProLeuIleLeuValAspSer	200
Db	600	ATAGGAACGAGGTTTACAAACCGTGACCGAGTCAACTCGTCTTTTAACTCTGTTGCTGCTCG	659
Qy	201	GlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSer	220
Db	660	CAGGACAACGAGTGTCTAGTCCACGCGCTTATGGCTTCCGCTGAAGCGCTGCAGAGC	719
Qy	221	SerAsnLeuThrIleAlaGluAlaLeuValLysGlnIleGlyPheLeuAlaValSerGln	240
Db	720	AGCAACCTTGACTCTAGCGGAGGCTCTCGTTAAGCAGATGCTTCTTGGCGGCTCTCAA	779
Qy	241	AlaGlyAlaMetArgLysValAlaThrTyrPheAlaGluAlaLeuAlaArgIleTyr	260

Db	780	GCCGGAGCCATGAGGAAAGTCGCCACGTACTTCGCCGAAGCTCTCGCGCGAGGATCTAC	839
Qy	261	ArgLeuSerProProGlnThrGlnIleAspHisSerLeuSerAspThrIleuGlnMetHis	280
Db	840	CGCCTCTCTCCGCGCAGACGCGAGATCGATCACTCTTTATCCGATATCTCTCCAGATGCAC	899
Qy	281	PheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeu	300
Db	900	TTCTACGAGACTTGCCTTACCTCAAGTTTCGCTCACTTACGGCGAATACAGGCGATCTC	959
Qy	301	GluAlaPheGluGlyLysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeu	320
Db	960	GAGGCTTTCGAAGGGAAGAGAGTCCAGTCATCGATTTCTCGATGAACCAAGGCGCT	1019
Qy	321	GlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyProProSerPheArg	340
Db	1020	CAGTGGCCCGCGCTTATGCAAGCCCTTGGCTTGAGGAAAGGAGTCTCTCCGAGTTTCAGG	1079
Qy	341	LeuThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCys	360
Db	1080	TTAACCGGAATTGCTCTCCCGCGCGATTACTCCGATCATCTCCATGAAGTTGATGT	1139
Qy	361	LysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAla	380
Db	1140	AACTTGGCTCAGCTCGCGAGGCGATTCACTCGAGTTTGAGTATCGTGGCTTTGTGCT	1199
Qy	381	AsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAla	400
Db	1200	AAATAGCTTAGCTGATCTTGATGCTTCGATGCTTTAGACCCGAGTGAACCGAAGCT	1259
Qy	401	ValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArgThrGlyGlyIleGlu	420
Db	1260	GTGCGCGTTAACTCTGTTTTCGAGCTCCCAAGCTCTTAGCCCGTACCGGTGGGATAGAG	1319
Qy	421	LysValPheGlyValValLysGlnIleLysProValIlePheThrValValGluGlnGlu	440
Db	1320	AAAGTCTCTGGCGTGTGAAACAGATTAAACCGGTGATTTTACCGTTGTTGAGCAAGAA	1379
Qy	441	SerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSer	460
Db	1380	TCGAATCATAAACCGTCCGGTTTTCTTAGACCGGTTTACTGAATCGCTGCATTTATTTCG	1439
Qy	461	ThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLysValMetSerGluVal	480
Db	1440	ACGTTGTTGATTCTTGGAGGTGCTCCGAGTAGCAAGATTAAGTTATGTCGGAAATT	1499
Qy	481	TyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyProAspArgValGluArg	500
Db	1500	TATTTAGGAAACAGATTTCGAATCTGGTGGCTTTCGAAGTCCGGACCGTGTGAGAGA	1559
Qy	501	HisGluThrLeuSerGlnTrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHis	520
Db	1560	CATGAGACGCTGAGTCAATGCTCAACCGGTTCCGGTTCGGTTCGGGTTTTCGCGCGGCGAT	1619
Qy	521	LeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGlu	540
Db	1620	CTCGGTCTTAACCGTTTAAGCAAGCAGTACGCTTTTGGCTTTGTTTAAATGAGGCGAA	1679
Qy	541	GlyTyrArgValGluLysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeu	560
Db	1680	GGTTATCGTGTGAGAGAATAATGGGTGTTTCATGTTGAGTTGGCACACTCGACCGCTC	1739
Qy	561	IleThrThrSerAlaTrpLysLeuSerAlaValHis	572
Db	1740	ATAACCACTCCGCTTGGAAAGCTCTCGCGGCTGCAC	1775

RESULT 2
AX081276
LOCUS
DEFINITION
AX081276
SEQUENCE 1 from Patent WO0109356.
ACCESSION
AX081276.1 GI:13170125

KEYWORDS	Brassica napus (rape)
SOURCE	Brassica napus
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE	1
AUTHORS	Renard,M., Delourme,R., Barret,P., Brunel,D., Proger,N. and Tanguy,X.
TITLE	Mutant gene of the gras family and plants with reduced development containing said mutant gene
JOURNAL	Patent: WO 0109356-A 1 08-FEB-2001;
FEATURES	INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR)
source	Location/Qualifiers
	1..1779
	/organism="Brassica napus"
	/mol_type="unassigned DNA"
	/db_xref="taxon:3708"
CDS	60..1778
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC33120.1"
	/db_xref="GI:13170126"
	/translation="MKRDILHQFGPNHGTSIAGSTSSPAVFCDKQMMVKEEDEL LGVLGKVRSEMAEVALKUEQLKLETMNGAQEDGLAHLATDTVHYNPABEYSLMDNML TELPSSPAATGSMALNPEINNNNNNSFTGTDLKAIPGNACRRSNQFAPAVDSNNK RLKPPSSPMVSTSPSPAGVIGTFTVTVTSTREPLILVDSQDNGVRLVHALMACAAV QSNLTLAEALVKIQIGFLAVSQAGAMRKVATYFAEALARRIYRLSPQOTQIDHLSLDT LQMHFETCYLFAFHTANOALAEAPGKKRVHVVDPSMNQGLWPALMQALALREG GPSPRLTGTGPPAUNDSDHLHEVGCKLAQAEAIHVEFYRGFVANSLADLDASMLE LRSETAEAVNSVFELHGLKRGTKERKVGVMQKIPKVFIVFGQESNHNPGPFLD RFTSEHLYSTLFDLSPESSQDKVMSEYVLGKICNLVACBGPQVERHETLSQWS NRFGSSGFAPAHLSGNAFKQASTLLALFNGEGGYRVEENNGCLMLSHTPLRTTTSAA KLSAVH"
ORIGIN	
Alignment Scores:	
Pred. No.:	1..61e-217
Score:	2946.00
Percent Similarity:	100.00%
Best Local Similarity:	99.83%
Query Match:	99.86%
DB:	6
	Length: 1779
	Matches: 571
	Conservative: 1
	Mismatches: 0
	Indels: 0
	Gaps: 0
US-110-030-194A-4 (1-572) x AX081276 (1-1779)	
Qy	1 MetIySArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db	60 ATGAAGAGGAGATCTTCATCATGTTCCAAAGGTCCCAACCCAGCGGACATCAATCGCGGTCT 119
Qy	21 SerThrSerProAlaValPheGlyLysAspLysMetMetValLysGluGluGlu 40
Db	120 TCCACTTCTCCCTCGCGGTGTTGGTAAAGCAAGATGATGATGTCGCAAGAAGAAGAA 179
Qy	41 AspAspGluLeuLeuGlyValLeuGlyTYrLysValArgSerSerGluMetalGluVal 60
Db	180 GACGACGAGCTTCTAGGAGCTTTGGGTTACAAGTTAGGCTTCGAGATGGCTGAGGTT 239
Qy	61 AlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGluAspGlyLeuAla 80
Db	240 GCGTTGAACCTCGAGCAGCTTGACAGCATGATGGGTTAACGCTCAAGAAGAACGCTT 299
Qy	81 HisLeuAlaThrAspThrValHisTyAsnProAlaGluLeuTySerTrpLeuAspAsn 100
Db	300 CACCTCGGACGGATATGTTCTATTACACCCCGCTGAGCTTTACTCGTGGCTTGATPAC 359
Qy	101 MetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAlaLeuAsnProGlu 120
Db	360 ATGCTCAGGAGCTTAACCCACCGCTGCAACGACCGCATCTAACGCTTTGMAACCCGAG 419
Qy	121 IleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLysAlaIleProGly 140
Db	420 ATTAATTAATTAATAAATCTGTTTTCACCGGAGCGCACTCAAGCGATTCCTGGA 479


```
Db 1059 AACCAAGCGATTCTCGAAGCTTTTGAAGGTAAAGAGAGATACACGTCATGATTGTTCTCG 1118
QY 316 MetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGly 335
Db 1119 ATGACCAAGCTCTTCATGCGCTTCGCTTATGCNAGCTCTTCGGCTTCGAGAGGAGGT 1178
QY 336 ProProSerPheArgLeuThrGlyLeuGlyProProAlaAlaAspAsnSerAspHisLeu 355
Db 1179 CCTCCAACTTTCGGTTAACCGAATTTGGTCCACCGCGCGGATTAATCTGATCATCTT 1238
QY 356 HisGluValGlyCysValLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyr 375
Db 1239 CATGAAGTGTGGTGTAAATTAGCTCAGCTTCGGAGCGGATTACGCTAGAAATTCGAATAC 1298
QY 376 ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395
Db 1299 CGTGGATTCTGCTTAACAGCTTAGCCGATCTCGATGCTTCGATGCTTAGACCG 1358
QY 396 SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuGlyArg 415
Db 1359 AGCGATACCGAAGCTGTTCGGGTGAACCTCTGTTTTGAGCTACATAAGCTCTTAGGTCTG 1418
QY 416 ThrGlyGlyLeuGlyValPheGlyValValLysGlnIleLysProValIlePheThr 435
Db 1419 CCCGGTGGGATAGAGAAAGTTCTCGCGGTGTGTGAACAGATTAAACCGGTGATTTTCACG 1478
QY 436 ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455
Db 1479 GTGCTTGAGCAAGATCGAACCATACCGACCGGTTTTCTTAGACCGGTTACTGAATCG 1538
QY 456 LeuHisTyrTrpSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys 475
Db 1539 TTACATTTATTTCCGACTCTGTTGATTCGTTGGAAGGAGTTCGGAATAGTCAAGACAA 1598
QY 476 ValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro 495
Db 1599 GTCATGTCTGAAGTTTACTTAGGAAACAGATTGTTGAATCTGTGGCTTGTGAAGTCTCT 1658
QY 496 AspArgValGluArgHisGluThrLeuSerGlnTrpSerAsnArgPheGlySerSerGly 515
Db 1659 GACAGAGTCGAGAGACACGAAAGCTTGAGTCATGGGAAACCGGTTGGTTCGTCGGT 1718
QY 516 PheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeu 535
Db 1719 TTAGCGCGGCACATCTTGGGCTTAACCGCTTTAAGCAAGCGAGTAGCTTTTGTCTGTG 1778
QY 536 PheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGlyCysLeuMetLeuSerTrp 555
Db 1779 TTTAATAGTGCCCAAGGTTATCGTGTGGAGGAGTAGTAATCGATGTTTGATGTTGGGTGG 1838
QY 556 HisThrArgProLeuThrThrSerAlaTrpLysLeuSer--AlaValHis 572
Db 1839 CACACTCGCCCACTCATACACTCGCTCGCTTGGAAACTCTCGACGGCGGCGCAC 1892

RESULT 4
AX507818
LOCUS 1764 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 2513 from Patent WO0216655.
ACCESSION AX507818
VERSION AX507818.1 GI:23389055
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
Harpur, J.F., Krops, J., Wang, X. and Zhu, T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 2513 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
```

```
(CH)
FEATURES
    source
    1..1764
    /organism="Arabidopsis thaliana"
    /mol_type="unassigned DNA"
    /db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 5.65e-176 Length: 1764
Score: 2407.00 Matches: 488
Percent Similarity: 86.03% Conservative: 23
Best Local Similarity: 82.15% Mismatches: 47
Query Match: 81.59% Indels: 36
DB: Gaps: 9
US-10-030-194A-4 (1-572) x AX507818 (1-1764)
QY 1 MetLysArgAspLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerIleAla 18
Db 1 ATGAAGAGAGATCATCACCATTCCAAAGGTTCGATTGTCCAAACACCGGACTTCTTCTTCA 60
QY 19 GlySerSerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGlu 38
Db 61 TCATCATCATCTCT-----AAAGATAGATGATGATGGTGAAGAAA 102
QY 39 GluGlu-----AspAspGluLeuGlyValLeuGlyTyrLysVal 52
Db 103 GAAGAAAGACGGTGGAGGTAAACATGGACGACGAGCTTCTCGCTGTTTAGGTTACAAAGTT 162
QY 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGly 72
Db 163 AGGTCTATCGGAGATCGCGAGGTGCTTTGAAACTCGAACAAATTAGAGACGATGATGAGT 222
QY 73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92
Db 223 AATGTTCAAGAGATGGTTTATCTCATCTCGACGCGATACTGTTTATATATATATCCGTCG 282
QY 93 GluLeuTyrSerTrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
Db 283 GAGCTTTATTTCTTGGCTTGATAATATGCTCTCTGAGCTTAATCTCTCTCTCTCTCCGCG 342
QY 113 GlySerAsnAlaLeuAsnProGluLysAsnAlaValCysArgArgSerAsnGlnPheAla 151
Db 343 AGTCTTAACGGTTTAGATCCGGTTCTCTCTCCCGGAGATTTGCGTTTTCGGGCTTCG 402
QY 132 GlyGlyAspLysAlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAla 151
Db 403 GATTATGACTTAAAGTCAATCCCGAAACCGGATT-----TATCAGTTTCCG 450
QY 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
Db 451 ---GCGATTGATTCTTCTGCTTCGTAATAATCAGAACAAAGGTTTGAATCATGTCG 507
QY 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182
Db 508 AGTCTCTGATTCTATGTTTATCATCGACTTCGACGGGTACGAGATTGGTGAGTCATAGGA 567
QY 183 ThrThrValThrThrValThr-----GluSerThrArgProLeu 195
Db 568 ACAGCGGTGACGCAACACCAACACGACGACGACGACGCGCGGGTGGTCACTGTTCTGTT 627
QY 196 IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
Db 628 ATCTCGTTGACTCGCAAGAGAACGTTGTTGTTTAGTCCACGCGCTTATGTTGTTGCA 687
QY 216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235
Db 688 GAAGCAATCCAGCAGAACAAATTTGACTAGCGGAAAGCTCTTGTGAAGCAAAATCGGATGC 747
QY 236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrThrPheAlaGluAlaLeu 255
Db 748 TTAGCTGTGTCTCAAGCCGAGGCTATGAGAAAAGTGGTACTTACTTTCGCCGAGCTTTA 807
```

QY	256	AlaArgArgIleTyrArgLeuSerProProGlnThrGlnIleAspHisSerLeuSerAsp	275
DB	808	GC CGCGGGGATCTACCGTCTCTTCGCGCGCAGAAATCAGATCATATGTCCTCTCCGAT	867
QY	276	ThrLeuGlnMethHisPheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAla	295
DB	868	ACTCTTCAGATGCACATTTTACGAGACTTGTCCTATCTTAAATTCGCTCACTTCACGGCG	927
QY	296	AsnGlnAlaIleLeuGluAlaPheGluGlyLysArgValHisValIleAspPheSer	315
DB	928	AACCAAGCGGATCTCGAAGCTTTTGAAGGTAAGCAAGAGAGTACACGTCATGATTTCTCG	987
QY	316	MetAsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGly	335
DB	988	ATGACCAAGGCTCTTCAATGGCCCTGCACATATGCAGCTCTTCGCGTTCGAGAAGAGGT	1047
QY	336	ProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAspHisSerAspHisLeu	355
DB	1048	CCTCCAACCTTCGCGTTAACCGGAATTTGGTCCACCGCGCGCGGATAATTCGATCATCTT	1107
QY	356	HisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyr	375
DB	1108	CATGAAGTTGGTTGTAATTAGCTCAGCTTGC CGAGCGCATTCACGTAGAAATTCGAATAC	1167
QY	376	ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro	395
DB	1168	CTGGATTCGTTGCTTACAGCTTAGCGGATCTCGATGCTTCGATGCTTAGCTTAGCCG	1227
QY	396	SerGluThrGluAlaValAlaValAnSerValPheGluLeuHisLysLeuLeuGlyArg	415
DB	1228	AGCGATACGGAAGCTGTTCGCGTGAATCTGTTTTGAGCTACATAAGCTCTTAGTCCGT	1287
QY	416	ThrGlyGlyIleGluLysValPheGlyValValLysGlnIleLysProValIlePheThr	435
DB	1288	CCCGGTGGGATAGAGAAAGTTCTCGCGGTGTGTAACAGATTAACCGGTGATTTTCACG	1347
QY	436	ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer	455
DB	1348	GTGTTGAGCGAGATCGAACCATACGAGACCGCGTTTTCTTAGACCGTTACTGATCG	1407
QY	456	LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys	475
DB	1408	TTACATATTATTTCGACTCTGTTTGTATTCGTTGGAAGGAGTTCGGAATAGTCAAGACAA	1467
QY	476	ValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro	495
DB	1468	GTCAATGCTGAAGCTTTACTTAGCGGAACAGATTGTAATCTGGTGGCTTGTGAAGTCC	1527
QY	496	AspArgValGluArgHisGluThrLeuSerGlnTyrPheSerAsnArgPheGlySerSerGly	515
DB	1528	GACAGAGTCGAGAGACACGNAAGTTGATGTCATGGGAAACCGGTTGGTTCGTCGGT	1587
QY	516	PheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeu	535
DB	1588	TTAGCGCGGCACATCTTGGGTCTAACCGGTTTAAAGCAAGCGAGTATGCTTTTGTCTGTG	1647
QY	536	PheAsnGlyGlyGluGlyTyrArgValGlyLysAsnAsnGlyCysLeuMetLeuSerTyr	555
DB	1648	TTTAAATAGTGGCCCAAGGTTATCTGTGTGGAGGAGTAATGATGTTTGTGTTGGGTGG	1707
QY	556	HisThrArgProLeuIleThrThrSerAlaTyrLysLeuSer	569
DB	1708	CACACTCGTCCATCTATTACCACCTCCGCTTGGAAACTCTCG	1749
RESULT	5		
LOCUS	BT010467		
DEFINITION	Arabidopsis thaliana At2g01570/P2I9.19 gene, complete cds.		
ACCESSION	BT010467		
VERSION	BT010467.1		
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		

ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 1764)
REFERENCE AUTHORS	Chen,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P., Dale,J.M., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R. Arabidopsis ORF clones 2 (bases 1 to 1764)
TITLE JOURNAL	Chen,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P., Dale,J.M., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R. Direct Submission Submitted (30-AUG-2003) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K. The Salk, Stanford, PCEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Chan, M.M., Chang,C.M., Dale,J.M., Hsuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera, C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and Ecker,J.R.
FEATURES	Location/Qualifiers 1. .1764 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosomes="2" /clone="U13937" /ecotypes="Columbia" /note="this clone is in PUNI 51" 1. .1764 /note="putative RCAL, gibberellin repsonse modulation protein," /codon_start=1 /product="At2g01570/P2I9.19" /protein_id="AAQ65090.1" /db_xref="GI:34365557" /translation="MKRDHQPQRLSNHGTSSSSSSISKDKMMVKKEEDGGNNMDD ELLAVLYKVRSEMAEVALKLELTMSNVQEDGLSHLADTVHYNSELYSWLDF MSELNPPPLPASSNGLDVLPSPICGFPASDYLKVPNGAIYQPPAIDSSSSNN QNKRLKSCSSPDSMTSITGQIGVITGTTVTITTTTAAGESTRSVLVDSQNG VRLVLMACAEAIQNNITLAEALVKQIGCLAVSQAGMRKVATYFASALARRIYL SPQNOIDHCLSDTLQMHFYETCPYLKFTANQAILEAFEGKKRVHVIDFSMNGL QWPAIMQALALREGGPTFRLTGIGFPAPDNSDHLHEVGCKLAQAEAIHVFEEYGF VANSIADLDASMLELRPSDTEAVANSVPEHLKLRPGGIEKVLKVPKIPVFTV VQESNNGHPVFLDRFTESLHYSTLFDLSLEGVPNSQDKVMSEVYLGKQICNLVACEG PDRVERHETLSQGNRFGSSGLAPAHLSNAFKQASMLLSVFNFSGGQYRVERSGCLM LGWHTRLPITTSWKLSTNAY"
CDS	Chen,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.
ORIGIN	

Alignment Scores:

Pred. No.: 5.65e-176 Length: 1764
 Score: 2407.00 Matches: 488
 Percent Similarity: 86.03% Conservative: 23
 Best Local Similarity: 82.15% Mismatches: 47
 Query Match: 81.59% Indels: 36
 DB: 8 Gaps: 9

US-10-030-194A-4 (1-572) x BT010467 (1-1764)

Qy 1 MetLysArgAspLeuHisGlnPheGlnGly-----ProHnHisGlyThrSerIleAla 18
 |||||
 Db 1 ATGAAGAGAGATCATCAACCAATTCCAAAGGTTCGATGTGCCAACACCGCGGACTTCTTCTCA 60
 |||||
 Qy 19 GlySerSerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGlu 38
 |||||
 Db 61 TCATCATCATCTCT-----AAAGATAAGATGATGATGGTGAAGAAA 102
 |||||
 Qy 39 GluGlu-----AspAspGluLeuLeuGlyValLeuGlyTyrLysVal 52
 |||||
 Db 103 GAAGAGACGGTGGAGGTAAACATGACGACGACGAGCTTCTCGCTGTTTTAGGTTACAAAGTT 162
 |||||
 Qy 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuThrMetMetGly 72
 |||||
 Db 163 AGGTCTATCGGAGATGGCGGAGGTGCTTTGAAACTCGAACAAATTAGACGACGATGATGAGT 222
 |||||
 Qy 73 AsnAlaGlnGluAspGlyLeuAlaHisAlaThrAspThrValHisTyrAsnProAla 92
 |||||
 Db 223 AATGTTCAAGAAATGTTTATCTCATCTCGACGAGATAGTGTTCATATAATATCGGTTCG 282
 |||||
 Qy 93 GluLeuTyrSerTyrLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
 |||||
 Db 283 GAGCTTTATTTCTGGCTTGATAATATGCTCTCTGAGCTTAATCTCTCTCTCTCTCTCGCGG 342
 |||||
 Qy 113 GlySerAsnAlaLeuAsnProGluIleAsnAsnAsnAsnAsnSerPhe----PheThr 131
 |||||
 Db 343 AGTCTTAACCGGTTAGATCCGGTTCCTCTCGCGGAGATTTGTGGTTTTCTCGCGCTTCG 402
 |||||
 Qy 132 GlyGlyAspLeuLysAlaIleProGlyAsnAlaValCysArgSerAsnGlnPheAla 151
 |||||
 Db 403 GATTATGACCTTAAGTCATCTCCGGAACCGCAT-----TATCAGTTTCCG 450
 |||||
 Qy 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
 |||||
 Db 451 ---GCGATTGATTTCTGCTCTCGTGAATAATCAGAACCAAGCGTTTGAATCATGCTCG 507
 |||||
 Qy 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182
 |||||
 Db 508 AGTCCTGATTTCTATGTTTACATCGACTTCAGCGGGTACGACGATTTGGTGGAGTCATAGGA 567
 |||||
 Qy 183 ThrThrValThrValThr-----GluSerThrArgProLeu 195
 |||||
 Db 568 ACAGCGGTGACCAACACCAACAGCAGACACAGCGCGCGGGTGGAGTCAATCTGTTCTGT 627
 |||||
 Qy 196 IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
 |||||
 Db 628 ATCTGGTTGACTCGCAAGAGACGGTGTTCGTTTGTAGTCCACGCGCTTATGGCTTGTGCA 687
 |||||
 Qy 216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235
 |||||
 Db 688 GAAGCAATCCAGCAGAACAAATTTGACTCTAGCGGAAGCTCTTGTGAAGCAAAATCGGATGC 747
 |||||
 Qy 236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrThrPheAlaGluAlaLeu 255
 |||||
 Db 748 TTAGCTGTGCTCAAGCGGAGCTATGAGAAAAGTGGCTACTTACTTCGCGCAAGCTTTA 807
 |||||
 Qy 256 AlaArgArgIleTyrArgLeuSerProProGlnThrGlnIleAspHisSerLeuSerAsp 275
 |||||
 Db 808 GCGCGCGGATCTACCGTCTCTCTCGCGCGCAAGAAATCAGATCATATGTTCTCTCTCGAT 867
 |||||
 Qy 276 ThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295
 |||||

Db 868 ACTCTTCAGATGCACCTTTTACGAGACTTGTGCTTAATCTTAATTCGCTCACTTCACGCGG 927
 Qy 296 AsnGlnAlaIleLeuGluAlaPheGluGlyLysLeuArgValHisValIleAspPheSer 315
 |||||
 Db 928 AACCAAGCGATTTCTCGAAGCTTTTGAAGGTAAAGAGAGATACACGTCATGATTTCTCG 987
 |||||
 Qy 316 MetAsnGlnGlyLeuGlnTyrProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGly 335
 |||||
 Db 988 ATGAACCAAGGCTCTCAATGGCTTGCACCTTATGCAAGCTTTCGCTTCGAGAGGAGGT 1047
 |||||
 Qy 336 ProProSerPheArgLeuThrGlyIleGlyProProAlaAlaAlaPheAsnSerAspHisLeu 355
 |||||
 Db 1048 CCTCCAACTTTCGGTTTAAACCGAATTTGTCACCGCGCGGATTAATCTGATCATCTT 1107
 |||||
 Qy 356 HisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyr 375
 |||||
 Db 1108 CATGAAGTTGGTTGTAATTAGCTCAGCTTTCGAGGCGGATTCACGTAGATTCGAATAC 1167
 |||||
 Qy 376 ArgGlyPheValAlaAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395
 |||||
 Db 1168 CGTGGATTCGTTGCTAACACGCTTAGCGATCTCGATGCTTCGATGCTTAGAGTTAGACCG 1227
 |||||
 Qy 396 SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuGlyArg 415
 |||||
 Db 1228 AGCGATACGGAAGCTGTTGCGGTAACTCTGTTTTTGGAGCTACATAAGCTCTTAGGTCGT 1287
 |||||
 Qy 416 ThrGlyGlyIleGluLysValPheGlyValValLysGlnIleLysProValIlePheThr 435
 |||||
 Db 1288 CCGGTTGGATAGAGAAAGTTCTCGCGGTTGTGAAACAGATTAAACCGGTGATTTTCACG 1347
 |||||
 Qy 436 ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455
 |||||
 Db 1348 GTGGTTGAGCAAGATCGAACCATACCGACCGCTTTCTTAGACCGGTTTACTGAATCG 1407
 |||||
 Qy 456 LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys 475
 |||||
 Db 1408 TTACATATTATTTCGACTCTGTTTGAATTCGTTGGAAGAGGTTCGGAATAGTCAAGACAAA 1467
 |||||
 Qy 476 ValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro 495
 |||||
 Db 1468 GTCATGCTGTAAGTTTACTTTAGGAAACAGATTGTGTAATCTGGTGGCTTGTGAAGTCT 1527
 |||||
 Qy 496 AspArgValGluArgHisGluThrLeuSerGlnTyrSerAsnArgPheGlySerSerGly 515
 |||||
 Db 1528 GACAGAGTCAGAGACACGAAACGTTGAGTCAATCGGAAACCGGTTTGGTTCGTCGGT 1587
 |||||
 Qy 516 PheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeu 535
 |||||
 Db 1588 TTAGCGCGCGCACATCTGGGTCTAACCGCTTTAAGCAAGCGAGTATGCTTTGTCTGTG 1647
 |||||
 Qy 536 PheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGlyCysLeuMetLeuSerTyr 555
 |||||
 Db 1648 TTTAATAGTGGCAAGGTTATCTGTGGAGAGAGTAATGGATGTTTGTATGTTGGGTGG 1707
 |||||
 Qy 556 HisThrArgProLeuIleThrThrSerAlaTyrLysLeuSer 569
 |||||
 Db 1708 CACACTCGTCCACTCATACCACCTCCGCTTGGAACCTCTCG 1749
 |||||

RESULT 6

AY054160 2252 bp mRNA linear PLN 30-SEP-2001
 LOCUS Arabidopsis thaliana At2g01570/F219.19 mRNA, complete cds.
 DEFINITION
 ACCESSION AY054160
 VERSION AY054160.1 GI:15809785
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 2252)
 Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
 Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,

Ishida, J., Jiang, P. X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.

Arabidopsis cDNA clones

Unpublished

REFERENCE AUTHORS

Percent Similarity: 86.03% Conservativity: 23
Best Local Similarity: 82.15% Mismatches: 47
Query Match: 81.59% Indels: 36
DB: 8 Gaps: 9
US-10-030-194A-4 (1-572) x AY054160 (1-2252)

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PDEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C. J., Koesema, E., Meyers, M. C., Shinn, P., Ban, J., Bowser, L., Dale, J. M., Goldsmith, A. D., Jiang, P. X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R. W., Theologis, A., and Ecker, J. R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers
1..2252
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAFL09-22-101 (R13937)"
/ecotype="Columbia"
1..189
190..1953
/note="putative RGAI, gibberellin response modulation protein"

5'UTR

CDS

/codon_start=1
/product="At2g01570/F2I9.19"
/protein_id="AAL06821.1"
/db_xref="GI:15809786"
translation="MKRDHFOGRLSNHCTSSSSISKDKMMVKKEDGGNNDD
ELIAGVIGKVRSEMAEVALKLEETMMSNVQEDGLSHLATVYHNPSELYSWLNN
MLSEINPPPLPASPNDLPVLPSPICGFPSPDYLKVPIGNAIYQFPALDSSSSNN
QNRKLSKSSPDMVTSSTGTQIGVIGTVTTTAAGESTRSLVDSQNG
VRLVHALMACAEAIQQNNLTALBAVKQIGLAVSQAGAMRVATYFAELARRIYRL
SPQNOQIDCLSDTLQMHVEYCTPYLKFAHTANOAILLEAFEGKRVHVIDFSGMNOGL
QWALMQALALRGGPTPFLTGTGPGAPDPSDHLHEVCCKLAQLAEALHVEYERGF
FANALDMSLELRPSDTEAVAVNSVFELHKLGRPGGIEKLVGVQVQIKPVIITV
VDESHNNGPVLDFRFTESLHYSTLFDLSLEGVPNSDKVMSEVYLGKQICNLVACEG
PDQVERHETLSQNGRFGSGLPAHLGSAFQAQSMLLSVFNSGGQYRVEESNGCLM
LGWHTPLITTSWKLSTRAY"
1954..2252

3'UTR

ORIGIN

Alignment Scores:
Pred. No.: 7,84e-176 Length: 2252
Score: 2407.00 Matches: 488

1 MetLysArgAspLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerLeuAla 18
190 ATGAAGAGAGATCATCAACCAATTCCAAGGTCGATTGTCACACCGGAGCTTCTCTTCA 249
19 GlySerSerThrSerSerProAlaValPheGlyLysAspLysMetMetMetValLysGlu 38
250 TCATCATCAATCTCT-----AAAGATAAGATGATGATGATGATGATGATGATGATGAT 291
39 GluGlu-----AspAspGluLeuLeuGlyValLeuGlyValLysVal 52
292 GAAGAGAGCGTGGAGGTTAACAATGACGACGACGACTCTCGCTGTTTAGTTCACAAAGT 351
53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGly 72
352 AGGTATCGGAGATGGCGAGGTGCTTTGAAACTCGAACAAATTAGACGATGATGATGATGAT 411
73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92
412 AATGTTCAAGAAAGATGTTTATCTCATCTCTCGGACGAGTACTGTTCAATTATATATCGGTCG 471
93 GluLeuTyrSerTrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
472 GAGCTTATCTCTGGCTTGATATATGCTCTCTGAGCTTAATCTCTCTCTCTCTCTCTCTCTCGGGCG 531
113 GlySerAsnAlaLeuAsnProGluLeuAsnAsnAsnAsnAsnAsnSerPhe---PheThr 131
532 AGTTCTAACGGTTAGATCCGGTCTCTCTCTCGCGGAGATTGTTGTTTTCGGGCTTCG 591
132 GlyGlyAspLeuLysAlaLeuProGlyAsnAlaValCysArgSerAsnGlnPheAla 151
592 GATTATGACCTTAAAGTCAATCCCGAAACGCGATT-----TATCAGTTTCGCG 639
152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
640 ---GCGATTGATCTCTGCTTCGTCGAATATATCAGACGAGTTCGTCGATGATGATGATGATGATGAT 696
167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValLysGly 182
697 AGTCCTGATTCTATGCTTACATCGACTTCGACGGGTACGACAGATTGGTGGAGTCATAGGA 756
183 ThrThrValThrThrValThr-----GluSerThrArgProLeu 195
757 ACGACGTCGACGACCAACACCGACGACGCGCGGGGTGAGTCAACTCGTCTCTGTT 816
196 IleLeuValAspSerGlnAspGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
817 ATCTGTTGATCTCGCAAGAACCGGTGTTGTTAGTTCACCGCGTTATGCTTCTGTCGA 876
216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnLysGlyPhe 235
877 GAAGCAATCCAGACAGAACAAATTGACTCTAGCGGAAGCTCTTGTGAAGCAAAATCGGATGC 936
236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrThrPheAlaGluAlaLeu 255
937 TTAGCTGTGCTCTCAAGCCGAGCATAGAGAAAAGTGGCTACTTACTTCCCGAAGCTTTA 996
256 AlaArgArgIleTyrArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAsp 275
997 GCGCGCGGATCTTACCGTCTCTCTCGCGCGAAGATCAGATCAGATCAGATCAGTCTCTCTCCGAT 1056
276 ThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295
1057 ACTCTTCAGATGCACCTTTTACGAGACTTGTCTCTTATCTTAAATTCGCTCACTTCACGGCG 1116
296 AsnGlnAlaIleLeuGluAlaPheGluGlyLysArgValHisValLysAspPheSer 315

```
Db 1117 AACCAAGCGATTCTCGAAGCTTTTGAAGTAAAGAGAGTACACGTCATTGATTTCTCG 1176
Qy 316 MetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGlyGly 335
Db 1177 ATGAACCAAGGTTCTTCAATGGCTGCACATTATGCAAGCTCTTGGCTTCGAGAGGAGGT 1236
Qy 336 ProProSerPheArgLeuThrGlyLeuGlyProProAlaAlaAspAsnSerAspHisLeu 355
Db 1237 COTCAACTTTCCGGTTAAACCGAATTGTCACCGCGCGCGGATTAATTTCTGATCATCTT 1296
Qy 356 HisGluValGlyCysLeuLeuAlaGlnLeuAlaGluAlaLeuHisValGluPheGluTrp 375
Db 1297 CATGAAGTTGGTTGTAATTTAGCTCAGCTTCGGAGGCGAATTCAGTAGAATTCGAATAC 1356
Qy 376 ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395
Db 1357 CGTGGATTCTGTTAAACAGCTTACCGGATCTCGATCTCGATCTTGGATCTTAGACCG 1416
Qy 396 SerGluThrGluAlaValAlaValAlaValAlaValAlaValAlaValAlaValAlaVal 415
Db 1417 AGCGATACCGAAGCTGTTGCGGTGAACCTCTGTTTTGAGCTACATAAGCTCTTAGGTCGT 1476
Qy 416 ThrGlyGlyIleGluLeuValPheGlyValValValGlnIleLeuProValIlePheThr 435
Db 1477 CCCGTGGGATAGAGAAAGTTCTCGCGTGTGAAACAGATTAAACCGGTGATTTTTCACG 1536
Qy 436 ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455
Db 1537 GTGGTTGAGCAAGATCGAACCATACCGACCGGTTTCTTAGACCGGTTTACTGAATCG 1596
Qy 456 LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys 475
Db 1597 TTACATTATTATTCGACTCTGTTGATTCGTTGGAAGGAGTTCCGATAGTCAAGACAAA 1656
Qy 476 ValMetSerGluValTyrLeuGlyGlnIleCysAsnLeuValAlaCysGluGlyPro 495
Db 1657 GTCATGTCGTAAGTTTACTTAGGAAACAGATTGTAATCTGGTGGCTGTGGAAGTCCT 1716
Qy 496 AspArgValGluArgHisGluThrLeuSerGlnTrpSerAsnArgPheGlySerSerGly 515
Db 1717 GACAGGTCGAGAGACACGAACTGTGATCAATCGGGAAACCGGTTTGGTTCGTCGGT 1776
Qy 516 PheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeu 535
Db 1777 TTAGCGCGGCACATCTTGGGTCTAACCGCTTAAAGCAAGCAGTATGCTTTTGTCTGTG 1836
Qy 536 PheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGlyCysLeuMetLeuSerTrp 555
Db 1837 TTTAATAGTGGCAAGGTTATCGTGTGGAGGAGATAATGGAATGTTGATGTTGGTGTG 1896
Qy 556 HisThrArgProLeuIleThrThrSerAlaTrpLysLeuSer 569
Db 1897 CACACTGTCCTCATTTACCACTTCGCTTGGAAACTCTCG 1938

RESULT 7
AC005560/c
LOCUS 103125 bp DNA linear PLN 11-MAR-2002
DEFINITION Arabidopsis thaliana chromosome 2 clone F219 map rga, complete
sequence.
ACCESSION AC005560
VERSION AC005560.3 GI:20197398
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 103125)
AUTHORS Rounsley, S.D., Lin, X., Kaul, S., Shea, T.P., Fujii, C.Y., Mason, T.M.,
Shen, M., Romling, C.M., Fraser, C.M., Somerville, C.R. and Venter, J.C.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 103125)
AUTHORS Lin, X.
```

```

TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
REFERENCE Medical Center Dr., Rockville, MD 20850, USA
AUTHORS 3 (bases 1 to 103125)
TOWN, C.D. and Kaul, S.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
COMMENT Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org
FEATURES
    Location/Qualifiers
        source
            1..103125
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /cultivar="Columbia"
            /db_xref="taxon:3702"
            /chromosome="2"
            /map="rga"
            /clone="F219"
        misc_feature
            1..10993
            /note="overlap with BAC clone F10A8
            (AC006200:83037..94029)."
            41..697
            /gene="At2g01390"
            /note="synonym: F219.1; predicted by genescan"
            <41..-697
            /gene="At2g01390"
            41..697
            /gene="At2g01390"
            /codon_start=1
            /product="hypothetical protein"
            /protein_id="AAC67316.1"
            /db_xref="GI:3785969"
            /translation="MKQNLVAVDILLNQMRDRIKLSDFVVSALITNCDRCRTGAS
            LAFYSLKMGHLKXSLALIGNPLRSNLPKVIYVVKVKAQHSGLGCGQAMLIH
            RLGRPRLAADVDFLLPDQKGVAAYTALMDVYISAGSEKAMKILREMRERIMP
            SLGYDVLSSLEKTSDFQKEVALLRKEKSLVASARFENVHVEDKICDLLFATNLL
            "
            complement(787..815)
            /rpt_family="At_rich"
            830..1787
            /gene="At2g01400"
            /note="synonym: F219.2; predicted by genescan; supported by
            full length cDNA: Ceres: 2422"
            join(830..994,1320..1787)
            /gene="At2g01400"
            join(875..994,1320..1640)
            /gene="At2g01400"
            /codon_start=1
            /product="expressed protein"
            /protein_id="AAC67317.1"
            /db_xref="GI:3785970"
            /translation="MARSCSIWTPVLISLSPVTGESPKISRRRVILATISIGSPQILLE
            AKIPEKQLOVLDSKDVSRNTMLYLAAGFLLGNGFLNGEAAEARVGRKENRKALEK
            LRAKASEPNKSGNQKIEKEKEVEVFPPLPPLVVEANLLQ"
            complement(1769..2932)
            /gene="At2g01410"
            /note="synonym: F219.3; predicted by genefinder"
            complement(<1769..>2932)
            /gene="At2g01410"
            complement(1769..2932)
            /gene="At2g01410"
            /codon_start=1
            /product="hypothetical protein"
            /protein_id="AAC67318.1"
            /db_xref="GI:3785971"
            /translation="MKANMSSSSSTRYSRFPFPLPAVLFLGLLI PSSADNRHVINF
            RSPGLYPEGLTWDPDRQHFVLSHSRTIHSVSDAGVLETILSLDLPENSTILGLAV
            DSTNRLLACQLSPPLPPFSALYSRLSGRRVFLSLPSLPDGGDDIARDVANDV
            AVDFKGNAYVTNSAKNFIWKVDRCGAASIFSKSPLFNSQPVAAADAFSDRCGLNGIV
            YISGYLAVQSNSTGKVPKVEDSGNARLVLLNGDLIAADGMRTRRRDCTVMVVSQKK
            LWLLKSQDSNSEGVIYDEIDLDIGFPTAVTACGRDLIVLYGRVWEGIMSGSSYKEEG
            AREWFGIEVWSEKEGGDKIWLIVLYLIGFYAFCFWRFQMKLLTNMDKKIT"
            complement(6662..9796)
```


Db	83562	AGTTCTAAACGGTTTAGATCCGGTTCTCTCTCGCCGGAGATTGTTGGTTTTCCTCCGGCTTCG	83503	...
Qy	132	GlyGlyAspLeuLeuValLeuProGlyAsnAlaValCysArgSerAsnGlnPheAla	151	...
Db	83502	GATTATGACCTTAAAGTCATTCCCGAAGCGGATT	83455	...TATCATGTTTCGG
Qy	152	PheAlaValAspSerSerSer	166	...AsnLysArgLeuLysProSerSer
Db	83454	---GCGATTGATCTCTCGTTCGTCGAATAATCAGAACCAAGCGTTTGAAATCATGCTCG	83398	...
Qy	167	SerProAspSerMetValThrSerProSerPro	182	...AlaGlyValIleGly
Db	83397	AGTCCTGATTCTATGGTTACATCGACTTCGACGGGTACGACGATTGGTGGAGTCATAGGA	83338	...
Qy	183	ThrThrValThrThrValThr	195	...GluSerThrArgProLeu
Db	83337	ACGACGGTGACACACACCACGACCAACGACGCGCGGGTGAGTCAACTGCTTCTGTT	83278	...
Qy	196	IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla	215	...
Db	83277	ATCTGGTTGACTCGCAAGAACGGTGTGTTTTAGTCCACGCGCTTATGGCTTGTGCA	83218	...
Qy	216	GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe	235	...
Db	83217	GAAGCAATCCAGCAGAACAAATTTGACTCTAGCGGAAGCTCTTGTGAAGCAATCGGATGC	83158	...
Qy	236	LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaAlaThrTyrPheAlaGluAlaLeu	255	...
Db	83157	TTAGCTGTGTCTCAAGCCGAGCTATGAGAAAGTGCTACTTACTTCCCGAGCTTTA	83098	...
Qy	256	AlaArgArgIleTyrArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAsp	275	...
Db	83097	GCGCGCGGATCTACCGTCTCTCTCCGCGCAGAAATCAGATCGATCATTTGCTCTCCGAT	83038	...
Qy	276	ThrLeuGlnMetHisPheTyrGluThrCysProTyrLeuLysPheAlaHisPheThrAla	295	...
Db	83037	ACTCTTCAGATGCATTTTACGAGACTGTGCTTATCTTAAATTCGCTCACTTTCACGGCG	82978	...
Qy	296	AsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgValHisValIleAspPheSer	315	...
Db	82977	AACCAAGCGATTCTCGAAGCTTTTGAAGGTAAAGAGAGATACAGTCATTAATTTCTCG	82918	...
Qy	316	MetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGlyGly	335	...
Db	82917	ATGAACCAAGGTCTTCAATGGCTCGCACTTATGCAAGCTCTTTCGCTTCGAGAAGAGGT	82858	...
Qy	336	ProProSerPheArgIleThrGlyIleGlyProProAlaAlaAspAsnSerAspHisLeu	355	...
Db	82857	CCCTCCAACCTTCGCGTTAACCGGAATGGTCCACCGCGCCGGATAAATTCGATCATCTT	82798	...
Qy	356	HisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyr	375	...
Db	82797	CATGAAGTTGGTTGTAATTAGCTCAGCTTCGCGAGCGCATTCAGTAGAAATTCGAATAC	82738	...
Qy	376	ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro	395	...
Db	82737	CGTGGATTCTGCTAACAGCTTAGCCGATCTCGATGCTTCGATGCTTGAGCTTAGACCG	82678	...
Qy	396	SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArg	415	...
Db	82677	AGCGATACGGAAGCTGTTCGGGTGAACCTCTGTTTTTGGAGTCACATAAGCTCTTAGGTGCT	82618	...
Qy	416	ThrGlyGlyIleGluLysValPheGlyValValLysGlnIleLysProValIlePheThr	435	...
Db	82617	CCCGGTGGGATAGAGAAAGTTCTCCGCGGTGTGAAACAGATTAAACCCGGTGATTTTCACG	82558	...
Qy	436	ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer	455	...
Db	82557	GTGGTTGAGCAAGATCGAACCATACGACCGGTTTTCTTAGACCGGTTTACTGAAATCG	82498	...
Qy	456	LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys	475	...

82497	TTACATTATTATTCGACTCTGTTGATTCGTTGGAGAGGATTCGGAATAGTCACAGAAA	82438
Qy	476 ValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro	495
Db	82437 GTCATGTCGAAGTTTACGGAACAAGATTTGTAATCTGCTGGCTTGTGAAGGTCT	82378
Qy	496 AspArgValGluIuArgHisGluThrLeuSerGlnTyrSerAsnArgPheGlySerSerGly	515
Db	82377 GACAGAGTCGAGAGACACGAAACGTTTGAGTCAATGGGAAACCGGTTTGGTTGCTCGCGGT	82318
Qy	516 PheAlaProAlaHisLeuGlySerIleAlaPheLeuGlnAlaSerThrLeuLeuAlaLeu	535
Db	82317 TTAGCGCGGCGCATCTTGGGTCTACCGGTTTAGCAAGCAGATGCTTTTGTGTGG	82258
Qy	536 PheAsnGlyGlyGluGlyTyrArgValGluLysAsnAsnGlnCysLeuMetLeuSerTyr	555
Db	82257 TTTTAATAGTGGCAAGGTTATCGTGTGGAGGAGTAAATGATGTTTGTATGTTGGGTGG	82198
Qy	556 HisThrArgProLeuIleThrThrSerAlaTyrLysLeuSer	569
Db	82197 CACACTCGTCCACTCATACCACCTCCGCTCGGTAAGTCTCG	82156
RESULT 8		
AY052239		
LOCUS	2216 bp mRNA linear PLN 09-SEP-2001	
DEFINITION	Arabidopsis thaliana At2g01570/F219.19 mRNA, complete cds.	
ACCESSION	AY052239	
VERSION	AY052239.1 GI:15529229	
KEYWORDS	FLJ CDNA.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eutroids II; Brassicales; Brassicaceae; Arabidopsis.	
AUTHORS	1 (bases 1 to 2216) Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowers,L., Carninci,P., Date,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.	
TITLE	Arabidopsis cDNA clones	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 2216) Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowers,L., Carninci,P., Date,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.	
AUTHORS	Submitted Direct Submission Submitted (16-AUG-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA	
TITLE	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA' : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.	
JOURNAL	The Salk, Stanford, PCEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowers,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.	
COMMENT		

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers
1..2216
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAF106-78-H02 (R13605)"
/ecotype="Columbia"
1..136
137..1900
/note="putative RGA1, gibberellin response modulation protein"
/codon_start=1
/product="At2g01570/F219.19"
/protein_id="AAK97709.1"
/db_xref="GI:15529230"
/translation="MKRDHQPQGLSNHGTSSSSSSISKDMMVMKKEEDGGGNMDD
ELLAVLGYKRSSENAVALKLEQLETMNSNVQEDGLSHLATDTVHYNPSELYISWLDN
MLSEINPPPLPSSNGLDPLVPSPEICGFPASDYDLKVPIGNALIVYFPADIDSSSSNN
QNRKLSGSSPDSSMTSTGTQIGVIGTCTTTTAAAGSTRSDILVDQENG
VRLVHALMACAEAIQNNLTAELVKOIGLAVSQAGAKRVATYFAEALARRIYRL
SPPNQIDHCLSDTLQMHFVETCPVLFKAFHTAOAILEAPEGKRVHVIDFSSMNGRL
QNPALMQALAREGGPPTFLTGTPAPDSDHLEHVGCKLAQLAEALHVEPFYRGF
VANSIADIASMLEPUSDTEAVAVSVFELKGLRGEGLEKVLGVVKKQIKPVFTV
VEQSNHNGPVLPDRFTSLHYSTLPDSLEGVPSNDKVMSEVYLGKQICNLVACEG
PDRVERHTLSQGNRFGSGGLAPHLGSAFNAKQASMLLSVFNSSQGYRVESNGCLM
LGWHTRPITTSNAWKLSLTAAY"
1901..2216

3'UTR ORIGIN

Alignment Scores:

Pred. No.: 1.86e-175 Length: 2216
Score: 2402.00 Matches: 488
Percent Similarity: 85.8% Conservative: 22
Best Local Similarity: 82.1% Mismatches: 48
Query Match: 81.4% Indels: 36
DB: 8 Gaps: 9

US-10-030-194A-4 (1-572) x AY052239 (1-2216)

Qy 1 MetLysArgAspLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerileAla 18
Db 137 ATGAAGAGAGATCATCACCACCAATTCGAAGTGCATGTTCCACACCGGGGACTTCTTCTTCA 196
Qy 19 GlySerThrSerSerProAlaValPheGlyLysAspLysMetMetMetValLysGlu 38
Db 197 TCATCATCATCTCT-----AAAGATAGATGATGATGTTGGTGAAGAAA 238
Qy 39 GluGlu-----AspAspGluLeuLeuGlyValLeuGlyTyrLysVal 52
Db 239 GAAGAAGACGGTGGAGGTAACATGCACGACGAGCTTCTCGCTGTTTAGTTACAAAGTT 298
Qy 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGly 72
Db 299 AGGTATCGGAGATGGCGAGGTTCCTTTGAACCTCGAACATTAAGACGATGATGAGT 358
Qy 73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92
Db 359 AATGTTCAAGAGATGGTTTATCTCATCTCGCGACGGATACCTGTTTCATTATATTCGTCG 418
Qy 93 GluLeuTyrSerTrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
Db 419 GAGCTTTATCTTGGCTTGATATATGCTCTCTGAGCTTAATCTCTCTCTCTCTCTCTCT 478
Qy 113 GlySerAsnAlaLeuAsnProGluLeuAsnAsnAsnAsnAsnSerPhe---PheThr 131
Db 479 AGTTCTAACGGTTTATAGATCCGGTTCTCTCTCGCGGAGATTTGTGGTTTTCGGGCTTCG 538
Qy 132 GlyGlyAspLeuLysAlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAla 151

Db 539 GATTATGACCTTAAAGTCATCTCCGGAACCGGATT-----TATCAGTTTCCG 586
Qy 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
Db 587 ---GCGATTGATTCTTCGTCTTCGTAATAATCAGAACAGCGTTTGAATCATGCTCG 643
Qy 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182
Db 644 AGTCTGATTCTATGGTTACATTCGACTTCAACGGGTACGAGATTGGTGGAGTCAATAGA 703
Qy 183 ThrThrValThrThrValThr-----GluSerThrArgProLeu 195
Db 704 ACGACGGTGACGACCAACCCACGACGACGCGCGCGGTGAGTCAACTCGTTCTGTAT 763
Qy 196 IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
Db 764 ATCTCTGGTTGATCTCGCAAGAGACGGTGTTCGTGTAGTCCACGGCTTATGGCTTGTGCA 823
Qy 216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235
Db 824 GAAGCATTCGACGACCAATTTGACTCTAGCGGAGCTCTTGTGAACCAATCGATGTC 883
Qy 236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrTyrPheAlaGluAlaLeu 255
Db 884 TTAGCTGTGTCTCAAGCGGAGCTATGAGAAAAGTGGCTACTTACTTCCGCCGAAGCTTTA 943
Qy 256 AlaArgArgIleTyrArgLeuSerProGlnThrGlnIleAspHisSerLeuSerAsp 275
Db 944 CGCGCGCGGATCTACCGTCTCTCCGCGGAGCAATCAGATCGATCATGTGCTCTCCCAT 1003
Qy 276 ThrLeuGlnMetHisPheThrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295
Db 1004 ACTCTTCAGATGCATTTTACGAGACTGTCTTATCTTAAATTCGCTCACCTTCACGGCG 1063
Qy 296 AsnGlnAlaIleLeuGluAlaPheGluGlyLysLysArgValHisValIleAspPheSer 315
Db 1064 AACCAAGCGATTCTCGAAGCTTTTGAAGGTAAGAAGAGAGTACACGTCATTCATCTT 1123
Qy 316 MetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGly 335
Db 1124 ATGAACCAAGTCTTCAATGGCTGCACTTATGCAAGCTCTTGGCTTCGAGAGAGAGGT 1183
Qy 336 ProProSerPheArgLeuThrGlyGlyProAlaAlaAspAsnSerAspHisLeu 355
Db 1184 CTCTCAACTTTCGGTTAACCGGAATGGTCCACCGCGCGGATAATTCATGATCATCTT 1243
Qy 356 HisGluValGlyCysLysLeuAlaGlnLeuAlaGluAlaIleHisValGluPheGluTyr 375
Db 1244 CATGAAGTTGGTTGTAATAATAGCTCAGCTTCGAGGCGGATTCAGCTAGAATTCGAATAC 1303
Qy 376 ArgGlyPheValAlaAsnSerLeuAlaAspLeuAspAlaSerMetLeuGluLeuArgPro 395
Db 1304 CGTGGATTTCGTTGCTTAACAGCTTAGCCGATCTCGATGCTTCGATGCTTGAGCTTAGACCG 1363
Qy 396 SerGluThrGluAlaValAlaValAsnSerValPheGluLeuHisLysLeuLeuGlyArg 415
Db 1364 AGCGATACGGAAGCTGTTGCGGTGAACCTCTGTTTGTAGCTACATAAGCTCTTAGTCTG 1423
Qy 416 ThrGlyGlyIleGluLysValPheGlyValValLysGlnIleLysProValIlePheThr 435
Db 1424 CCCGGTGGGATAGAGAAAGTTCTCGCGCTTGTGAACAGATTAACACCGGTGATTTTCAG 1483
Qy 436 ValValGluGlnGluSerAsnHisAsnGlyProValPheLeuAspArgPheThrGluSer 455
Db 1484 GTGGTTGACGAAGAAATCGAACCATTAACGACCGGTTTCTTAGACCGGTTTACTGAATCG 1543
Qy 456 LeuHisTyrTyrSerThrLeuPheAspSerLeuGluGlyAlaProSerSerGlnAspLys 475
Db 1544 TTACATTATATTGACCTCTGTTGATTTCGTTGGAGGAGTTCGGAATAGTCAAGACAA 1603
Qy 476 ValMetSerGluValTyrLeuGlyLysGlnIleCysAsnLeuValAlaCysGluGlyPro 495

```

Db 1504 GTCATGCTGAAGTTTACTTAGGGAACACAGATTGTAATCTGGTGGCTTGTGAAGGTCCT 1663
Qy 496 AspArgValGluArgHisGluThrLeuSerGlnTrpSerAsnArgPheGlySerSerGly 515
Db 1664 GACAGAGTCGAGACACAGAAACGTTGAGTCAATGGGGAACCGGTTTGGTTGCTCCGGT 1723
Qy 516 PheAlaProAlaHisLeuGlySerAsnAlaPheLysGlnAlaSerThrLeuLeuAlaLeu 535
Db 1724 TTACGCGCGGCACATCTTGGCTCTTAACGCGTTTAAGCAAGCGAGTATGCTTTTCTCTGTG 1783
Qy 536 PheAsnGlyGlyGluGlyTyrArgValGluLysAsnGlyCysLeuMetLeuSerTrp 555
Db 1784 TTTTAATAGTGGCAAGGTATCTGTCGGAGAGAGTATGATGTTTGATGTTGGTGGTGG 1843
Qy 556 HisThrArgProLeuIleThrSerAlaTrpLysLeuSer 569
Db 1844 CACACTCGTCCACTCAATTACCACCTCGCTTGGAAACTCTCG 1885

RESULT 9
ATY15194 2201 bp mRNA linear PLN 01-NOV-1997
LOCUS Arabidopsis thaliana mRNA for GRS protein.
DEFINITION Y15194
ACCESSION Y15194
VERSION Y15194.1 GI:2569939
KEYWORDS GRS protein.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
Peng, J., Carol, P., Richards, D.E., King, K.E., Cowling, R.J.,
Murphy, G.P. and Harberd, N.P.
The Arabidopsis GAI gene defines a signalling pathway that
negatively regulates gibberellin responses
Genes Dev.
2 (bases 1 to 2201)
Harberd, N.P.
Direct Submission
Submitted (22-OCT-1997) N.P. Harberd, John Innes Centre, Molecular
Genetics, Colney Lane, Norwich, NR4 7UJ, UK
FEATURES
Location/Qualifiers
1..2201
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
144..1907
/codon_start=1
/product="GRS protein"
/protein_id="CAA75493.1"
/db_xref="GI:2569940"
/db_xref="TrEMBL:O23725"
/translation="MKKDHQFQRLNHSHTSSSSISKDKMMVKBEDGGNNMDD
SLAVLGVYKRSMSAEVALLEQETMNSNVQEDGLSHLAADTVHNPSELYSWLDN
MELNPLPASPSSNDPLVPSPEICGFPASDYDLEVI PCNATYQPPAIDSSSSNN
QNKELKSCSPDMVTSTGTCTGIGVITVTVTVTAAAESTRSVILVDSOENG
VRLVHALMACAEALQNNLTLAEALVKQICLAUSQAGAMRKVATYFAEALARIYEL
SPQNDQLHDSLDLQHFHTCYLKFAPHTANQALIEAFEGKRVHVHIDFSNNQGL
QWPAQMOLALAREGGPTFTLTGIGPPADNSDHLHVGCCLAQLAEAVHVEFYRGF
VANSALDMLSLRPSDEAVAVSVFELHLRPGGTIEKVLGVYVKQI KPIVFTV
VEQSNHNGPVFLDRFTESLHYSTLFDLSLEGVPSQDKVMSVYLKQICNLVACSG
PDRVERHETLSQWNRFGSSGLAPAHLGSAFNAKQASMLLSVFNSSGQYRVEESNGCLM
LGWTRPLIITTSAWKLSAAAY"

```

ORIGIN

Alignment Scores:

Pred. No.:	4,48e-175	Length:	2201
Score:	2397.00	Matches:	485
Percent Similarity:	85.86%	Conservative:	25
Best Local Similarity:	81.65%	Mismatches:	48
Query Match:	81.25%	Indels:	36

```

DB: 8 Gaps: 9
US-10-030-194a-4 (1-572) x ATY15194 (1-2201)
Qy 1 MetLysArgAspLeuHisGlnPheGlnGly-----ProAsnHisGlyThrSerIleAla 18
Db 144 ATGAGAGAGATCATCATCAATTCAGAGTCGATTGTTCCACACCGGGACTTCTTCTCT 203
Qy 19 GlySerSerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGlu 38
Db 204 TCATCATCAATTTCT-----AAAGATAAGATGATGATGATGATAAAAAA 245
Qy 39 GluGlu-----AspAspGluLeuLeuGlyValLeuGlyTyrIleVal 52
Db 246 GAAGAACACGGTGGAGGTAAACATGGACGATGAGCTTCTCGCTGTTTAGGTACAAAGTT 305
Qy 53 ArgSerSerGluMetAlaGluValAlaLeuLysLeuGluLeuGluThrMetMetGly 72
Db 306 AGGTATCGGAGATGGCGAGGTGCTTTGAACTCGAACATTTAGACGATGATGATGATGAT 365
Qy 73 AsnAlaGlnGluAspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAla 92
Db 366 AATGTTCAAGAAAGATGTTTATCTCATCTCGCGCGGATACGTGTTCAATAAATCCGTCG 425
Qy 93 GluLeuTyrSerTrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThr 112
Db 426 GAGCTTTATTTCTGGCTTGATAATATGCTCTCTGAGCTTAATCTCTCTCTCTCTCTCT 485
Qy 113 GlySerAsnAlaLeuAsnProGluIleAsnAsnAsnAsnAsnSerPhe---PheThr 131
Db 486 AGTTCTAACGGTTAGATCCGGTCTCTCTCGCGGAGATTTGTGTTTTCGGCTTCG 545
Qy 132 GlyGlyAspLeuLysAlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAla 151
Db 546 GATTATGACCTTGAAGTCATTTCCCGGAAACGCGATT-----TATCAGTTTCCG 593
Qy 152 PheAlaValAspSerSerSer-----AsnLysArgLeuLysProSerSer 166
Db 594 ---GCGATTGATCTTCGCTCTCGTCAATAATCAGAACAAAGCGTTTGAATCATGTCG 650
Qy 167 SerProAspSerMetValThrSerProSerPro-----AlaGlyValIleGly 182
Db 651 AGTCTGATTCAATGGTTTACATCGACCTCGACGGGTACGAGATTGGTGGAGTTCATAGGA 710
Qy 183 ThrThrValThrThrValThr-----GluSerThrArgProLeu 195
Db 711 ACGACGCTGACACACACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 770
Qy 196 IleLeuValAspSerGlnAspAsnGlyValArgLeuValHisAlaLeuMetAlaCysAla 215
Db 771 ATCTGCTGTGACTCGCAAGAGACGCGTGTGTTTAGTACACGCGCTTATGCGCTTGTCGA 830
Qy 216 GluAlaValGlnSerSerAsnLeuThrLeuAlaGluAlaLeuValLysGlnIleGlyPhe 235
Db 831 GAAGCAATCCAGCAGAAATAATTTGACTCTACGCGAAGCTCTTGAGACCAATTCGATGCG 890
Qy 236 LeuAlaValSerGlnAlaGlyAlaMetArgLysValAlaThrTyrPheAlaGluAlaLeu 255
Db 891 TTAGCTGTGTCTCAAGCGGAGCTATCAGAAAAGTGGCTACTTACTTCTCGCCGAAAGCTTTA 950
Qy 256 AlaArgArgIleTyrArgLeuSerProProGlnThrGlnIleAspHisSerLeuSerAsp 275
Db 951 GCGCGCGGATCTACCGCTCTCTCTCGCGCGCAGAAATCAGATCGATCATTTGTCTCTCCGAT 1010
Qy 276 ThrLeuGlnMetHisPheThrGluThrCysProTyrLeuLysPheAlaHisPheThrAla 295
Db 1011 ACTCTTCAGATGACACTTTTACGAGACTTGCTTATCTTAAATTCGCTCACTTCACGCGG 1070
Qy 296 AsnGlnAlaIleLeuGluAlaPheGluGlyLysArgValHisValIleAspPheSer 315
Db 1071 AACCAAGGATCTCGAAGCTTTTGAAGGTAAAGAGAGATACAGTCATTTGATTTCTCG 1130
Qy 316 MetAsnGlnGlyLeuGlnTrpProAlaLeuMetGlnAlaLeuAlaLeuArgGluGly 335

```


Qy	286	ProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGly	305
Db	742	CCTTATCTCAAGTTCGCTCACTTCACGGCGAATCAAGCGATTCTCGAAGCTTTTCAAGGG	801
Qy	306	LysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeu	325
Db	802	AAGAAAGAGTTCATGTCATTGATTCCTATAGAGTCAAGGCTTCAATGGCCGCGCTT	861
Qy	326	MetGlnAlaLeuAlaLeuArgGluGlyGlyProSerPheArgLeuThrGlyIleGly	345
Db	862	ATGCAAGGCTCTGCGCTTCGACCTGGTGGTCTCTCTGTTTCCGGTTAACCGGAATTCGT	921
Qy	346	ProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeu	365
Db	922	CCACCGGCACCGATAATTCGATTATCTCATGAAGTTGGGTGAAGCTGGCTCATTTA	981
Qy	366	AlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAsp	385
Db	982	GCTGAGGGGATTTCAGTTGAGTTTGAGTACAGAGATTGTGGCTAACACTTTAGCTGAT	1041
Qy	386	LeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAlaValAlaValaAsnSer	405
Db	1042	CTTGATGCTTCGATGCTTGAGCTTAGACCAAGTAGATTGATCTGTTGGGTTAACTCT	1101
Qy	406	ValPheGluLeuHisLysLeuLeuGlyArgThrGlyGlyIleGluLysValPheGlyVal	425
Db	1102	GTTTTTCGAGCTTCACAAAGCTCTTGGGACGACCTGGTCGATCGATAAGGTTCTTGGTGTG	1161
Qy	426	ValLysGlnIleLysProValIlePheThrValValGluGlnGluSerAsnHisGly	445
Db	1162	GTGAATCAGATTAAACCGGAGATTTTCACTGTGGTTGACGAGGAATCGAACCATTAATAGT	1221
Qy	446	ProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSer	465
Db	1222	CGATTTCCTTAGATCGGTTTACTGAGTCGTTGCAATTAATCTCAGCGTTGTTGACTCG	1281
Qy	466	LeuGluGlyAlaProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGln	485
Db	1282	TTGGAAGGTGTACCGAGTGTCAAGACAAAGGTCATGTCGGAGTTTACTTGGGTAAACAG	1341
Qy	486	IleCysAsnLeuValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSer	505
Db	1342	ATCTCAACAGTTGTGGCTTGTGATGGACCTCGACCGATTGACGCGTCATGAAACGTTGAGT	1401
Qy	506	GlnTrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAla	525
Db	1402	CAGTGGAGAACCGGTTCGGGTCTGCTGGGTTCGGGCTGCACATATTGGTTTCGANTCGG	1461
Qy	526	PheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGlu	545
Db	1462	TTTAAGCAAGCGAGTATGCTTTTGGCTCTGTTCAACGGCGGTGAGGGTTATCGGGTGGAG	1521
Qy	546	LysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAla	565
Db	1522	GAGAGTCAGCGGTGTCTCATGTTGGGTTGGCACACACACCGCTCATAGCCACCTCGCT	1581
Qy	566	TrpLysLeuSer	569
Db	1582	TGGAACACTCTCC	1593
RESULT 11			
LOCUS	AY142002	1602 bp mRNA linear	PLN 02-SEP-2002
DEFINITION	Arabidopsis thaliana Atlg14920/F10B6_15 mRNA, complete cds.		
ACCESSION	AY142002		
VERSION	AY142002.1	GI:22655347	
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE			
AUTHORS	1 (bases 1 to 1602)		
	Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A. and Ecker, J.R.		
TITLE	Arabidopsis ORF clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1602)		
AUTHORS	Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A. and Ecker, J.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-AUG-2002) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
	The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.		
	Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.		
FEATURES	Location/Qualifiers		
source	1..1602		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="1"		
	/clone="U14047"		
	/ecotype="Columbia"		
	/note="This clone is in pUNI 51"		
CDS	1..1602		
	/note="signal response protein (GAI)"		
	/codon_start=1		
	/product="Atlg14920/F10B6_15"		
	/protein_id="AA098266.1"		
	/db_xref="GI:22655348"		
	/translation="MK3DHHHHQDKTMMNEEDCGMDELLAVLGKVRSSMA DVAKLEQLEVMNSNVQEDLSQLATETVHYNPAELYTWLDSMLTDLTPSSNARYDL KAI PGDAI LNQFAIDSASSNSQGGGDTTNNKRLKCSNGVETTTAETSRHVVLV DSQNGVRLVHALLACAEVQKENTLVAEALVKQIGFLAVSQIGAMRKVATYFABALA RRI VRLSPQSQPIDHSLDTLQMFETCTPYLKFPAHTANQAIIGAFQKRVHVIDF SMSGLOWPALMOALALRPGGPVFRLTGTPAPDNFDYLHEVCGLAHIAEAHVE PEYFGVANTLADLADSLMRPSEISAVNSVFLKHLGRGCAIDKVLGVNQIK PEITFVQESNHSNPFIDRFTESLHYSTLPSLEGVPSQDKMSEVILGKQICN VVACGPDVERHETLQGNRRNRFSGFAAAHIGSNAPKQASMLLALFNFGGEGYVEE SDGCLMLGWHTRLIATSAWKLSTN"		
ORIGIN			
Alignment Scores:	3.27e-153	Length:	1602
Pred. No.:	2111.00	Matches:	428
Score:			

Toriumi,M., Yanada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cdna clones
2 (bases 1 to 2097)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yanada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (28-AUG-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, POEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yanada,K., Yamamura, Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.
Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) Contributed equally to this work as PIs.
Location/Qualifiers
1. .2097
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL09-10-N02(R14047)"
/ecotype="Columbia"
1. .140
141. .1742
/note="signal response protein (GAI)"
/codon_start=1
/product="Atg14920/F1086_15"
/protein_id="AAL25607.1"
/db_xref="GI:16648834"
translational="MKRDHHHHQDKTMMNEEDDNGMDGLAVLGKVRSSSEA
DVAQKLEQLEVMVSNVQEDLSQLATETVHNPAELTLDLSMLTDLNPPSSNRRVLDL
KAIPGAILNQFADISSASSNGGGDTYNNRKLKCSNGVETTTATASRHRVLDL
DSQNGVRLVHALLACAAVOKENLTVAEALVKQIGFVLSQIGAMRKVATYFAEALA
RTYRLSPSSDTHSLSDTLQWHFYETCPYLKFAHFTANQALIEAFQGKKRVHVIDF
SMSQGLQWPAALMADPGGVPFRLTIGTPAPDNFDYLVHGVCKLHAEIAHVEIF
PEYRGVANTLADLASMLERLSEISVANSVPELHKLGRPGAIDKVLGVVNOIK
PELFTVYQESNNSPFLDRFTIESLHYITLFDSEGVSGQDKVMSEVYLKQKICN
VVAQCDGPDVRVERHETLSQWNRFGSFAAAHIGSNAPKQASMLLALFNGGEGYRVSEE
SDCLMLGWHTRELIATSAWLKSTN"
1743. .2097
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 4,69e-153 Length: 2097
Score: 2111.00 Matches: 428
Percent Similarity: 79.28% Conservative: 35
Best Local Similarity: 73.29% Mismatches: 53
Query Match: 71.56% Indels: 68
DB: 8 Gaps: 8
US-10-030-194A-4 (1-572) x AY058194 (1-2097)

US-10-030-194A-4 (1-572) x AY058194 (1-2097)

Qy 346 ProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLeuLeuAlaGlnLeu 365
Db 1062 CCACCGGCACCGATAATTTTCGATTATCTTCATGAGTTGGGTGTAAGTCGCTCATTTA 1121
Qy 366 AlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAsp 385
Db 1122 GCTGAGCGGATTCACGTTGAGTTGAGTACAGAGGATTTGGCTTAACACTTTAGCTGAT 1181
Qy 386 LeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAlaValAlaValaAsnSer 405
Db 1182 CTTGATGCTTCGATGCTTGAGCTTAGACCAAGTAGAGTTGAATCTGTTCGCGTTAACTCT 1241
Qy 406 ValPheGluLeuHisLeuLeuGlyArgThrGlyGlyIleGluLeuValPheGlyVal 425
Db 1242 GTTTCGAGCTTCACAGCTCTTGGGACGACTGCTGGTCGATCGATAAGTTCTTGGTGTG 1301
Qy 426 ValLysGlnIleLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGly 445
Db 1302 GTGAATCAGATTAAACCGGAGATTTTCACTGTGTTGAGCAGGAATCGAACCATATAGT 1361
Qy 446 ProValPheLeuAspArgPheThrCluSerLeuHisTyrTyrSerThrLeuPheAsnSer 465
Db 1362 CCGATTTTCTTAGATCGGTTTACTGAGTCGTTGCAATTAATTACTGACGCTTGTGACTCG 1421
Qy 466 LeuGluGlyAlaProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGln 485
Db 1422 TTGGNAGGTACCGAGTGGTCNAGACAGGTCTATGTCGAGGTTTACTTGGGTAACAG 1481
Qy 486 IleCysAsnLeuValAlaCysGluGlyProPheArgValGluArgHisGluThrLeuSer 505
Db 1482 ATCTCAACGTTGTGGCTTGTGATGGACTGACCTGACCGAGTTGAGCGTCATGAACGTTGAGT 1541
Qy 506 GlnTrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAla 525
Db 1542 CAGTGGAGAACCGGTTCCGGTCTCGGTTGCGGTCGCACATATTGGTTGGAATCGC 1601
Qy 526 PheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGluTyrArgValGlu 545
Db 1602 TTTAAGCAAGCGAGTATGCTTTTGGCTCTGTTCAACGCGGTGAGGTTATCGGTTGAG 1661
Qy 546 LysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeuIleThrSerAla 565
Db 1662 GAGAGTACGGCTGTCTCATGTTGGGTTGGCACACACACCGGCTCATAGCCACCTCGCT 1721
Qy 566 TrpLysLeuSer 569
Db 1722 TGGAAACTCTCC 1733
RESULT 13
AC006917 132699 bp DNA linear PLN 28-JUN-2000
LOCUS Genomic sequence for Arabidopsis thaliana BAC F1086 from chromosome
DEFINITION I, complete sequence.
ACCESSION AC006917
VERSION AC006917.6 GI:4757662
KEYWORDS HTG
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 132699)
REFERENCE
AUTHORS Chao,Q., Shinn,P., Dunn,P., Buehler,E., Kahn,S., Kim,C., Walker,M.,
Williams,S., Altafi,H., Araujo,R., Conn,L., Conway,A.B.,
Gonzalez,A., Hansen,N.F., Huizar,D., Kremenetskaia,I., Lenz,C.,
Li,J., Liu,S., Luros,S., Rowley,D., Schwartz,J., Toriumi,M.,
Vysotskaia,V., Yu,G., Davis,R.W., Federspiel,N.A., Theologis,A. and
Ecker,J.R.
Genomic sequence for Arabidopsis thaliana BAC F1086 from chromosome
TITLE I
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 132699)
AUTHORS Ecker,J.R.

TITLE
JOURNAL Direct Submission
Submitted (25-FEB-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
3 (bases 1 to 132699)
Ecker,J.R.
REFERENCE
AUTHORS Direct Submission
TITLE Arabidopsis thaliana Genome Center,
JOURNAL Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
4 (bases 1 to 132699)
Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q.,
Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,B., Chin,C.,
Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N.,
Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J.,
Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H.,
Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M.,
Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
Direct Submission
Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
On May 7, 1999 this sequence version replaced gi:4731042.
FEATURES
source
1. 132699
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
complement(join(144..278,360..491,576..673,814..1435,
1749..1867,2161..2310,2441..2565,2698..2742,2837..2902,
2979..3031,3117..3234,3314..3415,3498..3655,3741..3818,
3894..3985,4054..4104,4138..4420))
/note="similar to Na/H antiporter dbj|BAA89487.1"
/codon_start=1
/evidence=not experimental
/product="F1086.1"
/protein_id="AAF79251.1"
/db_xref="GI:8778242"
/translation="MTSIIIGNALPKPEKAISSYSAENDSSPVDVIFAGTSLVL
GTACRYLPNGTRVPYTVLVIGIFLQSLGILCLIKLYSVSIKILEKYGKTHNL
KGHGIRIWNGINPDLILAVLPVLLFESSFMDVHQIKRCMGOMVLLAGPVLIST
FCGLGALKLPFFYNNWMTKSLLLGGLGATDPVAVALLKELGASKKMTLLIDGESLM
NDGVSVVVOLFPPYKMGHNSDWGSIIFKLQNSFGVAGIGLAFGASVFWKFIEND
TMAOITVTLGASFYATQEWAGVSGILTVMLGMPFAAFARTAFAGDSHQSLHFW
EMAYIANTLVMSLGVIIAESVLSGOTISYKGVSSOFILRPDPFFTLFLEDVKGFN
SWFLFLYLVQSLRCVGVVLYPLLCSGYGLDWKESIILTWSGLRVQLFLCTG
GIVFTLVNGSTTQLLHLRLMDTLTKRKILEYTKFEMMKALKAFENIGDDEEL
GSADPTVIRHISLKDLEGQVNPDPGYSGLDPTNIMDIRFLNGLISQCVNNS
HCOTSVVPKVAIDSRLLPVSGVQAAYWEMLDGDRITCQTANVLQSVDEALDLVSTSS
LSDWRGLPVRVHPFNYYKFLQSKIIPHLVTHLIVERLESACYTSSAPRAHRIARQQ
LHI FLGNSNTASTVINESEVEGEAKQFLEDVRDSFPQVLSVLTROVTHVVLNHLNG
YIKMLEKVLLEGKEVSHLHDVQSDLKLLRHPPSLKLPNVDDLITSNPLLKDRSSP
RSLAIGBTDA"
join(5650..5801,6081..6121,6230..6819)
/note="unknown protein; similar to EST emb|F14394.1"
/codon_start=1
/evidence=not experimental
/product="F1086.2"
/protein_id="AAF79216.1"
/db_xref="GI:8778207"
/translation="MRTPTTILLVGAIFSGAGYVRSDSADSHRYKEGDTVPLVANKY
GPFNPSETVRYFDLPFCIPGEVKKEALGVNLGDLVSAFKLNPFRDKESEVYC
NKLKSEVQKFRKAEKDYVYFOMYDDLPIWFGIKVDKDIKSDPSFKPYFLYKHQ
FEILYKNDRVLEISARMDPHSYLVLTDEKEDVAEFMTVKWKEETPEPEKMEKYSMS
SSLPHHLEHWFHSIINSCVTLTGTLATILMRVLKNDFMK"
join(6981..7029,7140..7296,7392..7876,7961..8180)
/note="similar to transmembrane 9 superfamily member 2
gi|4759874; similar to dbj|AV442489.1, gb|A1996021.1,
emb|F14395.1"
/codon_start=1
/evidence=not experimental
/product="F1086.3"

Db 392 GACGATCTTTCTCAACTCGCTACTGAGACTGTTCACTATAATCCCGCGAGCTTTACACG 451
Qy 97 TrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAla 116
Db 452 TGGCTTGATCTCTATGCTCACCAGCACTTAATCTCGTGGTCT----- 493
Qy 117 LeuAsnProGluIleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLys 136
Db 494 ---AACGCGAGTAC-----GATCTTAA 514
Qy 137 AlaIleProGlyAsnAlaValCysArgSerAsnGlnPheAlaPheAlaValAspSer 156
Db 515 GCTATCCGGTGACGCGATCTC-----AATCAGTTGCGCTATCGATTCGCTCT 565
Qy 157 SerSer-----AsnLysArgLeuLysProSer 165
Db 566 TCGTCTAACCAAGCGCGGAGGAGATACGTATATACTACAACAACGCGTTGAATGCTCA 625
Qy 166 SerSerProAspSerMetValThrSerProSerProAlaGlyValIleGlyThrThrVal 185
Db 626 AAC-----GGCGTCGTGGAACCAACC----- 646
Qy 186 ThrThrValThrGluSerThrArgProLeuIleLeuValLeuValLeuValSerAsnLeuThrLeu 225
Db 647 ACAGCGAGCGGTGAGTCAACTCGCATGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 706
Qy 206 ArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeu 225
Db 707 CGTCTCGTTACGCGCTTTGCGCTTGGCTGAAGCTGTTTCAAGAGAGAACTGACTGTG 766
Qy 226 AlaGluAlaLeuValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArg 245
Db 767 GCGGAAGCTCTGGTGAAGCAATCGGATCTTAGCTGTTTCTCAAATCGGAGCTATGAGA 826
Qy 246 LysValAlaThrTyPheAlaGluAlaGluAlaArgArgIleTyArgLeuSerProPro 265
Db 827 AAAGTCGTACTTACTTCCGCAAGCTCTCGCGCGCGGATTTACCGTCTCTCCGCTG 886
Qy 266 GlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrgluThrCys 285
Db 887 CAGAGTCCAATCGACCACTCTCTCCGATACCTCTTCAGATGCACTTCTACGAGACTTGT 946
Qy 286 ProTyrgluLysPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGly 305
Db 947 CTTATCTCAAGTTCGCTCACTTCAAGCGCATCAAGCGATTTCTCAAGCTTTTCAAGG 1006
Qy 306 LysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeu 325
Db 1007 AAGAAAGAGTTTCATGTCATTTCTATGATCAAGGTCTTCAATGGCGCGGCTT 1066
Qy 326 MetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGly 345
Db 1067 ATGAGAGCTCTTGGCTTCGACCTCGCTGCTCTCTGTTTCCGGTTAAACCGAATGCT 1126
Qy 346 ProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeu 365
Db 1127 CCACCGGACCGGATAAATTCGATTATCTCATGAAGTTGGGTGAAGTGGCTCATTTA 1186
Qy 366 AlaGluAlaIleHisValGluPheGluTyArgGlyPheValAlaAsnSerLeuAlaAsp 385
Db 1187 GCTGAGGCGATTCAGTTGAGTTGAGTACAGAGATTTGGGCTAAACACTTACTGAT 1246
Qy 386 LeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSer 405
Db 1247 CTTGATGCTTCGATGCTTGAGCTTAGACCAAGTAGATTTGAATCTGTTGGGTTAACTCT 1306
Qy 406 ValPheGluLeuHisLysLeuLeuGlyArgThrGlyIleGluLysValPheGlyVal 425
Db 1307 GTTTTCGAGCTTCAACAGCTCTTGGGACGACCTGGTGGCATCGATAAGGTTCTTGGTGTG 1366
Qy 426 ValLysGlnIleLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGly 445
Db 1367 GTGAATCAGATTAAACCGGAGATTTTCACTGCTGGTTGAGCAGGAATCGAACCATATAGT 1426

Qy 446 ProValPheLeuAspArgPheThrGluSerLeuHisTyTySerThrLeuPheAspSer 465
Db 1427 CCGATTTTCTTAGATCGGTTTACTCAGTCGTTGCTGCTATTACTCGACGTTGTTGACTCG 1486
Qy 466 LeuGluGlyAlaProSerSerGlnAspLysValMetSerGluValTyLeuGlyLysGln 485
Db 1487 TTGGAGGTGTACCGAGTGGTCAAGCAAGGTCTGTCGAGGTTTACTTTGGTTAAACAG 1546
Qy 486 IleCysAsnLeuValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSer 505
Db 1547 ATCTGCAACGTTGTGGCTTGTGATGACCTGACCGAGTTGAGCGTCATGAACGTTGACT 1606
Qy 506 GlnTrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAla 525
Db 1607 CAGTCGAGAAACCGGTTCCGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1666
Qy 526 PheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyArgValGlu 545
Db 1667 TTTAAGCAAGCGAGTATGCTTTTGGCTCTGTTCAACGCGGTTGAGGTTTATCGGGTGGAG 1726
Qy 546 LysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAla 565
Db 1727 GAGAGTGACGCTCTCTCATGTTGGGTTGGCACACACGACCGCTCATAGCCACTCGGCT 1786
Qy 566 TrpLysLeuSer 569
Db 1787 TGGAAACTCTCC 1798
RESULT 15
LOCUS ARI174879 1964 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 1 from patent US 6307126.
ACCESSION ARI174879
VERSION ARI174879.1 GI:17915199
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1964)
AUTHORS Harber,N.P., Peng,J., Richards,D.E. and Carol,P.
TITLE Nucleic acid encoding GAI gene of Arabidopsis thaliana
JOURNAL Patent: US 6307126-A 1 23-OCT-2001;
FEATURES
source Location/Qualifiers
1..1964
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 4,696-153 Length: 1964
Score: 2110.50 Matches: 429
Percent Similarity: 79.45% Conservative: 35
Best Local Similarity: 73.46% Mismatches: 51
Query Match: 71.54% Indels: 69
DB: 6 Gaps: 9
US-10-030-194A-4 (1-572) x ARI174879 (1-1964)
Qy 1 MetLysArgAspLeuHisGlnPheGlnGlyProAsnHisGlyThrSerIleAlaGlySer 20
Db 209 ATGAAGAGAGATCATCATCAT-----CATCATCAAGATAAAG----- 244
Qy 21 SerThrSerSerProAlaValPheGlyLysAspLysMetMetValLysGluGlu 40
Db 245 -----AAGACTATGATGATGATGAAGAGAC 271
Qy 41 Asp-----AspGluLeuLeuGlyValLeuGlyTyTyLysValArgSerGlu 56
Db 272 GACGGTAACCGCATGATGAGCTTCTAGCTGTTCTTGGTTACCAAGGTTAGGTTCATCGGAA 331
Qy 57 MetAlaGluValAlaLeuLysLeuGluGlnLeuGluThrMetMetGlyAsnAlaGlnGlu 76
Db 577 GTGAATCAGATTAAACCGGAGATTTTCACTGCTGGTTGAGCAGGAATCGAACCATATAGT 1426

Db 332 ATGGCTGATGTTGGCTCAGAAACTCGAGCAGCTTGAAGTTATGATGTCTAATGTTCAAGAA 391
Qy 77 AspGlyLeuAlaHisLeuAlaThrAspThrValHisTyrAsnProAlaGluLeuTyrSer 96
Db 392 GACGATCTTCTCAACTCGCTACTGAGACTGTTCACTATAATCCGGCGGAGCTTTACAG 451
Qy 97 TrpLeuAspAsnMetLeuThrGluLeuAsnProProAlaAlaThrThrGlySerAsnAla 116
Db 452 TGGCTTGATCTATGCTCACCGACCTTAATCTCGTGTCT----- 493
Qy 117 LeuAsnProGluIleAsnAsnAsnAsnAsnSerPhePheThrGlyGlyAspLeuLys 136
Db 494 ---AACCGCGAGTAC-----GATCTTAAA 514
Qy 137 AlaIleProGlyAsnAlaValCysArgArgSerAsnGlnPheAlaPheAlaValAspSer 156
Db 515 GCTATTCCTCGGTGACCGCATCTC-----AATCAGTTCCGTATCGATTCCGCTTCT 565
Qy 157 SerSer-----AsnLysArgLeuLysProSer 165
Db 566 TCGCTTAACCAAGCGCGGAGGAGATACGTATATACTACAAACAAGCGGTTGAAATGCTCA 625
Qy 166 SerSerProAspSerMetValThrSerProSerProAlaGlyValIleGlyThrThrVal 185
Db 626 AAC-----GGCGTCGTGGAAACCAACC----- 646
Qy 186 ThrThrValThrGluSerThrArgProLeuIleLeuValAspSerGlnAspAsnGlyVal 205
Db 647 ACAGGAGCGGTGAGTCAACTCGGCATGTTGCTCGTTGACTCGCAGGAGAACGGTGTG 706
Qy 206 ArgLeuValHisAlaLeuMetAlaCysAlaGluAlaValGlnSerSerAsnLeuThrLeu 225
Db 707 CGTCTCGTTACCGCGCTTTGGCTTCGCTGAGCTGTTTCAGAGGAGAACTCGACTGTG 766
Qy 226 AlaGluAlaLeuValLysGlnIleGlyPheLeuAlaValSerGlnAlaGlyAlaMetArg 245
Db 767 CGCGAAGCTCTGGTGAAGCAAAATCGGATCTTAGCTGTTTCTCAAAATCGGAGCTATGAGA 826
Qy 246 LysValAlaThrTyrPheAlaGluAlaLeuAlaArgArgIleTyrArgLeuSerProPro 265
Db 827 AAAGTCGCTACTTACTTCGCGGAAGCTCTCGCGCGCGGATTTACCGTCTCTCTCCGTG 886
Qy 266 GlnThrGlnIleAspHisSerLeuSerAspThrLeuGlnMetHisPheTyrGluThrCys 285
Db 887 CAGAGTCCAATCGACCACTCTCTCTCCGATCTCTCAGATGCACCTTCTACGAGACTTGT 946
Qy 286 ProTyrLeuLysPheAlaHisPheThrAlaAsnGlnAlaIleLeuGluAlaPheGluGly 305
Db 947 CTTTATCTCAAGTTTCGCTCACTTCACGGCGGAATCAAGCGATTTCTCGAAGCTTTTCAAGGG 1006
Qy 306 LysLysArgValHisValIleAspPheSerMetAsnGlnGlyLeuGlnTrpProAlaLeu 325
Db 1007 AAGAAAGAGTTCAATGATTTCTCTAGATCAAGGTCTTCAATGGCGCGGCTT 1066
Qy 326 MetGlnAlaLeuAlaLeuArgGluGlyGlyProProSerPheArgLeuThrGlyIleGly 345
Db 1067 ATGAGAGCTCTTGGCTTCGACCTGGTGGTCTCTCTCTCTTTCGGTTAAACCGGAATTGGT 1126
Qy 346 ProProAlaAlaAspAsnSerAspHisLeuHisGluValGlyCysLysLeuAlaGlnLeu 365
Db 1127 CCACCGGACCGGATAATTCGATTATCTTCATGAAGTTGGGTGTAAGCTGGCTCATTTA 1186
Qy 366 AlaGluAlaIleHisValGluPheGluTyrArgGlyPheValAlaAsnSerLeuAlaAsp 385
Db 1187 GCTGAGGCGATTTCAGTTGAGTTGAGTACAGAGGATTTGGGCTTAACACITTAGCTGAT 1246
Qy 386 LeuAspAlaSerMetLeuGluLeuArgProSerGluThrGluAlaValAlaValAsnSer 405
Db 1247 CTTGATGCTTCGATGCTTGAGCTTAGACCAAGTGAAGATTGAATCTGTTGGGTTAACTCT 1306
Qy 406 ValPheGluLeuHisLysLeuLeuGlyArgThrGlyIleGluLysValPheGlyVal 425
Db 1307 GTTTTCGAGCTTCACAAGCTCTTCGAGACGACCTGGTGGCATCGATAAGGTTCTTGGTGTG 1366

Qy 426 ValLysGlnIleLysProValIlePheThrValValGluGlnGluSerAsnHisAsnGly 445
Db 1367 GTGAATCAGATTAAACCGGAGATTTTCACTCTGTGTTGAGCAGGAATCGAACATAATAGT 1426
Qy 446 ProValPheLeuAspArgPheThrGluSerLeuHisTyrTyrSerThrLeuPheAspSer 465
Db 1427 CCGATTTTCTAGATCGGTTTACTGAGTCGTTGCATTATTACTCGACGTTGTTTGACTCG 1486
Qy 466 LeuGluGlyAlaProSerSerGlnAspLysValMetSerGluValTyrLeuGlyLysGln 485
Db 1487 TTGAAGGTGTACCAGTGTGTCACAGACAGGTCAATGTCGGAGGTTTACTTGGGTAAACAG 1546
Qy 486 IleCysAsnLeuValAlaCysGluGlyProAspArgValGluArgHisGluThrLeuSer 505
Db 1547 ATCTGCAACGTTGTGGCTTGTGATGGACCTGACCGAGTTGAGCGTCATGAAACGTTGAGT 1606
Qy 506 GlnTrpSerAsnArgPheGlySerSerGlyPheAlaProAlaHisLeuGlySerAsnAla 525
Db 1607 CAGTGGAGGAACCGGTTTCGGGTCCTGCTGGGTTTCGGCTGCACATATTTGGTTCGAATGCG 1666
Qy 526 PheLysGlnAlaSerThrLeuLeuAlaLeuPheAsnGlyGlyGluGlyTyrArgValGlu 545
Db 1667 TTTAAGCAAGCAGATATGCTTTTGGCTCTGTTCAACGGCGGTGAGGTTATCGSGTGGAG 1726
Qy 546 LysAsnAsnGlyCysLeuMetLeuSerTrpHisThrArgProLeuIleThrThrSerAla 565
Db 1727 GAGAGTCAGCGCTCTCTCATGTTGGGTTGGCACACGACGCGCTCATAGCCACCTCGGCT 1786
Qy 566 TrpLysLeuSer 569
Db 1787 TGGAAACTCTCC 1798

Search completed: November 3, 2004, 09:34:50
Job time : 9566 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2004, 11:25:27 ; Search time 2090.5 Seconds
(without alignments)
135.727 Million cell updates/sec

Title: US-10-030-194A-5

Perfect score: 28

Sequence: 1 GYXVEE 6

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 5.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/p/US10030194/runat.01112004.184847.28955/app.query.fasta_1.398
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10030194 @CGN 1 1 7589 @runat.01112004.184847.28955 -NCPU=6 -ICPU=3
-NO MMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	27	96.4	200 6	C0713883
2	27	96.4	318 6	AX310498
3	27	96.4	330 6	I08256
4	27	96.4	387 6	I08257

5	27	96.4	388	11	G79421	G79421 S210P6193RA
6	27	96.4	400	6	CQ475389	CQ475389 Sequence
7	27	96.4	431	6	CQ505357	CQ505357 Sequence
8	27	96.4	479	6	AX435885	AX435885 Sequence
9	27	96.4	490	6	AX306056	AX306056 Sequence
c 10	27	96.4	498	6	AX400752	AX400752 Sequence
11	27	96.4	503	14	AY584590	AY584590 Nootka lu
12	27	96.4	535	6	AX432800	AX432800 Sequence
13	27	96.4	565	6	CQ527229	CQ527229 Sequence
c 14	27	96.4	582	11	BV074683	BV074683 S209P6108
c 15	27	96.4	622	11	G93830	G93830 S208P6399RE
c 16	27	96.4	623	11	G86206	G86206 S208P6399RE
c 17	27	96.4	650	9	AF399524	AF399524 Homo sapi
c 18	27	96.4	662	10	AY170535	AY170535 Mus muscu
c 19	27	96.4	681	11	BV042227	BV042227 S212P6005
c 20	27	96.4	708	6	BD214758	BD214758 Novel hum
c 21	27	96.4	732	1	AY102903	AY102903 Unculture
c 22	27	96.4	772	5	AY662689	AY662689 Gallus ga
c 23	27	96.4	849	6	AB346929	AB346929 Sequence
c 24	27	96.4	873	8	BT012980	BT012980 Lycopersi
c 25	27	96.4	926	5	BC071487	BC071487 Danio rer
c 26	27	96.4	963	6	AX241744	AX241744 Sequence
c 27	27	96.4	964	3	AK062617	AK062617 Oryza sat
c 28	27	96.4	986	3	AY267671	AY267671 Bigelowie
c 29	27	96.4	987	6	AX318263	AX318263 Sequence
c 30	27	96.4	987	6	AX448585	AX448585 Sequence
c 31	27	96.4	987	6	BD144456	BD144456 Novel G-p
c 32	27	96.4	1002	6	AX556236	AX556236 Sequence
c 33	27	96.4	1035	12	AY659169	AY659169 Synthetic
c 34	27	96.4	1061	5	CF523161	CF523161 Gallus ga
c 35	27	96.4	1064	8	AB059484	AB059484 Cucurbita
c 36	27	96.4	1065	6	AB346914	AB346914 Sequence
c 37	27	96.4	1068	10	AF441835	AF441835 Mus muscu
c 38	27	96.4	1115	9	AF489251	AF489251 Chiropete
c 39	27	96.4	1203	5	BX930000	BX930000 Gallus ga
c 40	27	96.4	1214	5	CF524283	CF524283 Gallus ga
c 41	27	96.4	1220	5	BX933929	BX933929 Gallus ga
c 42	27	96.4	1221	3	S55498	S55498 G alpha 4-G
c 43	27	96.4	1239	6	CQ645844	CQ645844 Sequence
c 44	27	96.4	1242	6	AX607019	AX607019 Sequence
c 45	27	96.4	1296	6	CQ821806	CQ821806 Sequence
c 46	27	96.4	1300	1	SAU19899	SAU19899 Streptococc
c 47	27	96.4	1301	1	S81109	S81109 citrate syn
c 48	27	96.4	1366	1	TC16SITAL	TC16SITAL
c 49	27	96.4	1387	6	AX646063	AX646063 Sequence
c 50	27	96.4	1387	9	AB065783	AB065783 Homo sapi
c 51	27	96.4	1398	8	AY034379	AY034379 Capsicum
c 52	27	96.4	1421	1	TC16SCZEC	TC16SCZEC
c 53	27	96.4	1428	1	AXU88041	AXU88041 Amaricoccus
c 54	27	96.4	1430	8	FSU23722	FSU23722 Fusarium so
c 55	27	96.4	1476	6	CQ729805	CQ729805 Sequence
c 56	27	96.4	1491	10	MAUGCC10	MAUGCC10
c 57	27	96.4	1497	6	AX078547	AX078547 Mesocricetu
c 58	27	96.4	1507	5	XELTRANS	XELTRANS
c 59	27	96.4	1529	1	SHU12238	SHU12238
c 60	27	96.4	1531	5	BC046071	BC046071 Danio rer
c 61	27	96.4	1586	8	SCCYTB8R	SCCYTB8R
c 62	27	96.4	1607	10	AF041083	AF041083 Rattus no
c 63	27	96.4	1657	5	AY292498	AY292498 Triopter
c 64	27	96.4	1659	6	AX621378	AX621378 Sequence
c 65	27	96.4	1806	5	AY292497	AY292497 Triopter
c 66	27	96.4	1835	5	AY292496	AY292496 Micromesi
c 67	27	96.4	1865	5	AY292490	AY292490 Boreogadu
c 68	27	96.4	1894	8	AFU277436	AFU277436 Aspergill
c 69	27	96.4	1896	6	AX079054	AX079054 Sequence
c 70	27	96.4	1932	5	AY292485	AY292485 Arctogadu
c 71	27	96.4	2000	6	AX595606	AX595606 Sequence
c 72	27	96.4	2000	6	AX819744	AX819744 Sequence
c 73	27	96.4	2006	6	AX830774	AX830774 Sequence
c 74	27	96.4	2036	6	AX927178	AX927178 Sequence
c 75	27	96.4	2104	6	AX834957	AX834957 Sequence
c 76	27	96.4	2104	6	AX097682	AX097682 Homo sapi
c 77	27	96.4	2262	6	CQ729806	CQ729806 Sequence

78	27	96.4	2289	6	AR072901	Sequence	AR072901	Sequence	c 151	27	96.4	36273	9	HUM7501	L78442 Homo sapien
79	27	96.4	2289	6	AX469427	Sequence	AX469427	Sequence	c 152	27	96.4	36731	9	AP006221	AP006221 Homo sapi
80	27	96.4	2289	6	BD080290	Isolation	BD080290	Isolation	c 153	27	96.4	36795	9	U88170	U88170 Caenorhabdi
81	27	96.4	2289	8	YSCSCS1A	L33931 Saccharomyc	L33931	Saccharomyc	c 154	27	96.4	36905	2	AL131062	AL131062 Homo sapi
82	27	96.4	2552	8	AY170625	AY170625 Penicilli	AY170625	Penicilli	c 155	27	96.4	37655	2	AL451184	AL451184 Human DNA
83	27	96.4	2591	14	AY277917	AY277917 Human rot	AY277917	Human rot	c 156	27	96.4	37701	3	AF000262	AF000262 Caenorhab
84	27	96.4	2591	14	AY277918	AY277918 Human rot	AY277918	Human rot	c 157	27	96.4	37763	2	AC130701	AC130701 Homo sapi
85	27	96.4	2591	14	AY277920	AY277920 Human rot	AY277920	Human rot	c 158	27	96.4	38561	2	AC008993	AC008993 Homo sapi
86	27	96.4	2640	8	AF347059	AF347059 Mycosphae	AF347059	Mycosphae	c 159	27	96.4	42041	3	U53181	U53181 Caenorhabdi
87	27	96.4	3018	8	YSCSUT1	L27744 Yeast SUT1	L27744	Yeast SUT1	c 160	27	96.4	43284	3	AL133228	AL133228 Human DNA
88	27	96.4	3202	1	AF157829	AF157829 Myxococcu	AF157829	Myxococcu	c 161	27	96.4	53889	8	FL4010	AC026234 Sequence
89	27	96.4	3645	6	C0490818	C0490818 Sequence	C0490818	Sequence	c 162	27	96.4	54253	9	AF176915	AF176915 Homo sapi
90	27	96.4	3645	6	C0492795	C0492795 Sequence	C0492795	Sequence	c 163	27	96.4	54526	2	AC087404	AC087404 Homo sapi
91	27	96.4	3645	6	C0496663	C0496663 Sequence	C0496663	Sequence	c 164	27	96.4	55790	8	AB025602	AB025602 Arabidops
92	27	96.4	3645	6	C0497698	C0497698 Sequence	C0497698	Sequence	c 165	27	96.4	58764	10	AL845276	AL845276 Mouse DNA
93	27	96.4	3900	1	AB020482	AB020482 Aetopyrum	AB020482	Aetopyrum	c 166	27	96.4	59641	10	D78344	D78344 Mouse DNA F
94	27	96.4	3930	1	BSPHRST	BSPHRST	BSPHRST		c 167	27	96.4	60797	2	AL355356	AL355356 Homo sapi
95	27	96.4	3962	6	AX451794	AX451794 Sequence	AX451794	Sequence	c 168	27	96.4	61467	2	AC087637	AC087637 Homo sapi
96	27	96.4	3962	9	AF273049	AF273049 Homo sapi	AF273049	Homo sapi	c 169	27	96.4	62576	9	AC108122	AC108122 Homo sapi
97	27	96.4	4075	9	AF534078	AF534078 Homo sapi	AF534078	Homo sapi	c 170	27	96.4	63655	2	AC073024	AC073024 Homo sapi
98	27	96.4	4102	8	YSCSTO1X	L07650 Saccharomyc	L07650	Saccharomyc	c 171	27	96.4	63655	2	AL603746	AL603746 Human DNA
99	27	96.4	4223	8	YSCGCR3	D10224 Saccharomyc	D10224	Saccharomyc	c 172	27	96.4	64848	9	AL603746	AL603746 Human DNA
100	27	96.4	4274	2	AC143094	AC143094	AC143094		c 173	27	96.4	65326	2	AL124815	AL124815 Mus muscu
101	27	96.4	4312	10	BC068136	BC068136 Mus muscu	BC068136	Mus muscu	c 174	27	96.4	65663	2	AC104955	AC104955 Homo sapi
102	27	96.4	4661	8	AF198116	AF198116 Emericell	AF198116	Emericell	c 175	27	96.4	65811	9	AL390069	AL390069 Human DNA
103	27	96.4	4874	3	BT004830	BT004830 Drosophil	BT004830	Drosophil	c 176	27	96.4	66004	3	CEC07A9	Z29094 Caenorhabdi
104	27	96.4	4874	3	BT004830	BT004830 Drosophil	BT004830	Drosophil	c 177	27	96.4	66719	4	BX649259	BX649259 Smnthops
105	27	96.4	5742	1	AE005589	AE005589 Escherich	AE005589	Escherich	c 178	27	96.4	67163	2	AC090294	AC090294 Homo sapi
106	27	96.4	5925	6	AX346478	AX346478 Sequence	AX346478	Sequence	c 179	27	96.4	67506	2	AC087630	AC087630 Homo sapi
107	27	96.4	5929	7	AX39255882	AX392558 Lactobaci	AX392558	Lactobaci	c 180	27	96.4	67506	2	AC105223	AC105223 Homo sapi
108	27	96.4	6258	1	AF179847	AF179847 Lactococc	AF179847	Lactococc	c 181	27	96.4	69921	2	AC105223	AC105223 Homo sapi
109	27	96.4	6354	9	HSW805727	HSW805727	HSW805727		c 182	27	96.4	70658	10	BX119970	BX119970 Mouse DNA
110	27	96.4	6571	6	AX337792	AX337792 Sequence	AX337792	Sequence	c 183	27	96.4	71215	2	AC026514	AC026514 Homo sapi
111	27	96.4	6571	9	HUMN2220	D83032 Homo sapien	D83032	Homo sapien	c 184	27	96.4	71784	2	AC141990	AC141990 Rattus no
112	27	96.4	6759	1	AE009785	AE009785 Pyrobacul	AE009785	Pyrobacul	c 185	27	96.4	72000	9	BS000595	BS000595 Pan trogl
113	27	96.4	7618	1	AE005734	AE005734 Caulobact	AE005734	Caulobact	c 186	27	96.4	72745	8	AC133337	AC133337 Oryza sat
114	27	96.4	7618	1	AE005734	AE005734 Caulobact	AE005734	Caulobact	c 187	27	96.4	73998	2	AC073187	AC073187 Homo sapi
115	27	96.4	7957	1	AF010951	AF010951 Methanosa	AF010951	Methanosa	c 188	27	96.4	75523	9	AL627233	AL627233 Human DNA
116	27	96.4	10091	1	AF135398	AF135398 Thermus b	AF135398	Thermus b	c 189	27	96.4	77448	8	AFASCI11	AL713629 Aspergill
117	27	96.4	10214	1	AE0111477	AE0111477 Leptospir	AE0111477	Leptospir	c 190	27	96.4	77849	2	AC067872	AC067872 Homo sapi
118	27	96.4	10353	1	AE013568	AE013568 Methanosa	AE013568	Methanosa	c 191	27	96.4	78449	2	AC141965	AC141965 Rattus no
119	27	96.4	10648	1	AE012510	AE012510 Xanthomon	AE012510	Xanthomon	c 192	27	96.4	78835	8	AC120528	AC120528 Oryza sat
120	27	96.4	10804	1	AE013495	AE013495 Methanosa	AE013495	Methanosa	c 193	27	96.4	78956	2	AC022808	AC022808 Homo sapi
121	27	96.4	10805	1	AE013048	AE013048 Thermoana	AE013048	Thermoana	c 194	27	96.4	80598	9	AL158191	AL158191 Human DNA
122	27	96.4	11255	1	AE012545	AE012545 Xanthomon	AE012545	Xanthomon	c 195	27	96.4	80598	9	AL158191	AL158191 Human DNA
123	27	96.4	11274	1	AE011482	AE011482 Leptospir	AE011482	Leptospir	c 196	27	96.4	81095	2	AC113319	Continuation (4 of
124	27	96.4	11508	1	AE015373	AE015373 Shigella	AE015373	Shigella	c 197	27	96.4	82939	9	AL135928	AL135928 Human DNA
125	27	96.4	11787	1	AE010308	AE010308 Methanopy	AE010308	Methanopy	c 198	27	96.4	83621	2	AC148612	Continuation (4 of
126	27	96.4	11990	1	AE010918	AE010918 Methanosa	AE010918	Methanosa	c 199	27	96.4	84232	2	AC148612	Continuation (4 of
127	27	96.4	12177	7	P29LATE2	M14782 Bacillus ph	M14782	Bacillus ph	c 200	27	96.4	85110	9	AL133331	AL133331 Human DNA
128	27	96.4	12495	1	AE004470	AE004470 Pseudomon	AE004470	Pseudomon	c 201	27	96.4	85532	1	AY090559	AY090559 Providenc
129	27	96.4	12558	1	AE010146	AE010146 Pyrococcu	AE010146	Pyrococcu	c 202	27	96.4	89009	1	BS275208	B. subtilis
130	27	96.4	13321	6	AX353886	AX353886 Sequence	AX353886	Sequence	c 203	27	96.4	89790	2	AC025596	AC025596 Homo sapi
131	27	96.4	13885	1	AE006326	AE006326 Lactococc	AE006326	Lactococc	c 204	27	96.4	89795	8	NC3H10	AL513442 Neurospor
132	27	96.4	15443	1	AE000748	AE000748 Aquifex a	AE000748	Aquifex a	c 205	27	96.4	89934	2	AC008867	AC008867 Homo sapi
133	27	96.4	16279	8	SC9553	Z48622 S.cerevisia	Z48622	S.cerevisia	c 206	27	96.4	89992	9	AL135794	AL135794 Human DNA
134	27	96.4	16448	1	AF337958	AF337958 Streptoco	AF337958	Streptoco	c 207	27	96.4	90064	9	AC116376	AC116376 Homo sapi
135	27	96.4	17000	6	AF337093	AF337093 Sequence	AF337093	Sequence	c 208	27	96.4	90175	2	AP0092391	AP0092391 Homo sapi
136	27	96.4	17276	1	AF163833	AF163833 Streptoco	AF163833	Streptoco	c 209	27	96.4	93033	2	AP000644	AP000644 Homo sapi
137	27	96.4	17596	1	AF355776	AF355776 Streptoco	AF355776	Streptoco	c 210	27	96.4	94483	1	AY055428	AY055428 Human DNA
138	27	96.4	17719	9	AY254046	AY254046 Homo sapi	AY254046	Homo sapi	c 211	27	96.4	95956	6	AX602206	AX602206 Sequence
139	27	96.4	18239	1	AF349539	AF349539 Streptoco	AF349539	Streptoco	c 212	27	96.4	96394	2	AC079260	AC079260 Homo sapi
140	27	96.4	19365	7	P2ACG	M11813 Bacterioph	M11813	Bacterioph	c 213	27	96.4	96598	6	AX695500	AX695500 Sequence
141	27	96.4	21882	10	AL606782	AL606782 Mouse DNA	AL606782	Mouse DNA	c 214	27	96.4	96715	9	AL161621	AL161621 Human DNA
142	27	96.4	21972	9	AL840643	AL840643 Human DNA	AL840643	Human DNA	c 215	27	96.4	97906	9	AL449043	AL449043 Human DNA
143	27	96.4	22800	1	AE014244	AE014244 Streptoco	AE014244	Streptoco	c 216	27	96.4	98062	9	HSU0776P7	HSU0776P7
144	27	96.4	25021	1	AB028896	AB028896 Streptoco	AB028896	Streptoco	c 217	27	96.4	99483	1	AY055428	AY055428 Vibrio ch
145	27	96.4	28291	3	CEF56H9	Z74473 Caenorhabdi	Z74473	Caenorhabdi	c 218	27	96.4	99575	9	AL365318	AL365318 Human DNA
146	27	96.4	29954	5	BX649440	BX649440 Zebrafish	BX649440	Zebrafish	c 219	27	96.4	100605	2	AC087157	AC087157 Mus muscu
147	27	96.4	33954	8	SC9905	Z46861 S.cerevisia	Z46861	S.cerevisia	c 220	27	96.4	101220	9	AC092372	AC092372 Homo sapi
148	27	96.4	35061	9	AC096569	AC096569 Homo sapi	AC096569	Homo sapi	c 221	27	96.4	102644	6	AX770900	AX770900 Sequence
149	27	96.4	35310	2	AC073804	Continuation (6 of	Continuation (6 of		c 222	27	96.4	106540	8	AC129091	AC129091 Medicago
150	27	96.4	35679	3	CBRG05D12	AC084487 Caenorhab	AC084487	Caenorhab	c 223	27	96.4	106771	2	AC135231	AC135231

C 224	27	96.4	109512	2	AC138251	Mus muscu	C 297	27	96.4	138403	2	AC138165	Bos tauru
C 225	27	96.4	110000	1	AE017225_36	Continuation (37 o	C 298	27	96.4	139054	8	AC130805	Medicago
C 226	27	96.4	110000	37	AE017225_37	Continuation (38 o	C 299	27	96.4	139056	8	AC139670	Medicago
C 227	27	96.4	110000	1	AE017334_36	Continuation (37 o	C 300	27	96.4	139086	2	AL1356127	Human DNA
C 228	27	96.4	110000	1	AE017334_37	Continuation (38 o	C 301	27	96.4	139908	2	AL1423368	Rattus no
C 229	27	96.4	110000	1	AE017355_37	Continuation (38 o	C 302	27	96.4	140907	9	AL162499	Human DNA
C 230	27	96.4	110000	1	EX571856_24	Continuation (25 o	C 303	27	96.4	140942	9	AC002525	Human DNA
C 231	27	96.4	110000	1	EX571857_23	Continuation (24 o	C 304	27	96.4	141492	2	AC093204	Homo sapi
C 232	27	96.4	110000	1	BX950851_18	Continuation (34 o	C 305	27	96.4	141924	8	AL138830	Human DNA
C 233	27	96.4	110000	1	CR543861_21	Continuation (12 o	C 306	27	96.4	142571	8	AP003316	Oryza sat
C 234	27	96.4	110000	1	U00096_37	Continuation (38 o	C 307	27	96.4	142596	2	AC140874	Homo sapi
C 235	27	96.4	110000	2	AC091370_0	Continuation (38 o	C 308	27	96.4	142596	1	AE016794	Homo sapi
C 236	27	96.4	110000	2	AC076768_1	Continuation (2 of	C 309	27	96.4	142565	1	AE016794	Pseudomon
C 237	27	96.4	110000	2	AC110933_2	Continuation (3 of	C 310	27	96.4	142737	8	AC027658	Oryza sat
C 238	27	96.4	110000	2	AC120804_3	Continuation (4 of	C 311	27	96.4	143079	2	AC021449	Homo sapi
C 239	27	96.4	110000	2	AC127890_2	Continuation (4 of	C 312	27	96.4	143236	2	AC069054	Homo sapi
C 240	27	96.4	110000	2	AC018724_2	Continuation (3 of	C 313	27	96.4	143516	9	AC083783	Homo sapi
C 241	27	96.4	110000	2	AC148014_0	Continuation (3 of	C 314	27	96.4	143615	9	AC104006	Homo sapi
C 242	27	96.4	110000	2	AC148612_1	Continuation (2 of	C 315	27	96.4	143895	8	CNS08CD6	Oryza sat
C 243	27	96.4	110000	2	AC148612_1	Continuation (2 of	C 316	27	96.4	144201	9	HS230119	Human DNA s
C 244	27	96.4	110000	2	AP006501_03	Continuation (4 of	C 317	27	96.4	144202	2	AC110179	Mus muscu
C 245	27	96.4	110000	2	BX284653_4	Continuation (5 of	C 318	27	96.4	144749	8	AC025912	Mus muscu
C 246	27	96.4	110000	2	EX901978_1	Continuation (2 of	C 319	27	96.4	145579	2	AC015737	Oryza sat
C 247	27	96.4	110000	2	AC073804_4	Continuation (5 of	C 320	27	96.4	145579	2	AC008531	Homo sapi
C 248	27	96.4	110000	8	CR380957_11	Continuation (12 o	C 321	27	96.4	145737	2	AC116898	Homo sapi
C 249	27	96.4	110000	8	CR382123_10	Continuation (11 o	C 322	27	96.4	145911	1	AP003014	Mesorhizo
C 250	27	96.4	110000	8	CR382127_05	Continuation (6 of	C 323	27	96.4	146949	2	AC1360003	Homo sapi
C 251	27	96.4	110185	9	AC079795_1	Continuation (6 of	C 324	27	96.4	147505	10	BX572085	Mouse DNA
C 252	27	96.4	110347	2	AC046193	AC079795 Homo sapi	C 325	27	96.4	148113	2	AL1391624	Homo sapi
C 253	27	96.4	110347	2	AC046193	AC046193 Homo sapi	C 326	27	96.4	148584	2	AC141682	Apis mell
C 254	27	96.4	110359	9	AC006158	AC006158 Homo sapi	C 327	27	96.4	148854	10	AC127557	Mus muscu
C 255	27	96.4	112581	9	AC006156	AC006156 Homo sapi	C 328	27	96.4	148887	8	AP003448	Oryza sat
C 256	27	96.4	112696	8	AC093225	AC093225 Homo sapi	C 329	27	96.4	148933	9	AC140232	Homo sapi
C 257	27	96.4	112721	8	AC016780	AC016780 Genomic S	C 330	27	96.4	148959	5	AC147203	Xenopus t
C 258	27	96.4	112866	8	AP005556	AP005556 Oryza sat	C 331	27	96.4	148978	9	HSAT9612	Homo sapi
C 259	27	96.4	115339	8	AP003940	AP003940 Oryza sat	C 332	27	96.4	149000	9	AC127540	Homo sapi
C 260	27	96.4	115469	9	AL356219	AL356219 Human DNA	C 333	27	96.4	149083	2	AC115528	Rattus no
C 261	27	96.4	115641	8	F8K4	AC004392 Arabidops	C 334	27	96.4	149138	5	BX548254	Zebrafish
C 262	27	96.4	115963	8	F28J9	AC007918 Arabidops	C 335	27	96.4	149195	9	AC116343	Homo sapi
C 263	27	96.4	117309	9	AC007274	AC007274 Homo sapi	C 336	27	96.4	149301	9	AC044849	Homo sapi
C 264	27	96.4	118092	2	AC147616	AC147616 Mus muscu	C 337	27	96.4	149495	2	AC106900	Homo sapi
C 265	27	96.4	118251	2	AC147180	AC147180 Phytophth	C 338	27	96.4	150321	2	AC097582	Rattus no
C 266	27	96.4	119389	9	AL450328	AL450328 Human DNA	C 339	27	96.4	150495	2	AC046148	Mus muscu
C 267	27	96.4	120851	2	AC016515	AC016515 Homo sapi	C 340	27	96.4	150595	9	AC146223	Pan trogl
C 268	27	96.4	121101	2	AC008034	AC008034 Homo sapi	C 341	27	96.4	150635	4	AC147165	Pan trogl
C 269	27	96.4	121258	9	AC010141	AC010141 Homo sapi	C 342	27	96.4	150771	9	AL392106	Human DNA
C 270	27	96.4	121741	5	EX649222	EX649222 Chicken D	C 343	27	96.4	151414	2	AC115523	Rattus no
C 271	27	96.4	121803	9	AC004835	AC004835 Homo sapi	C 344	27	96.4	151578	2	AP001074	Homo sapi
C 272	27	96.4	121949	9	AL157772	AL157772 Human DNA	C 345	27	96.4	151687	2	AL356979	Homo sapi
C 273	27	96.4	122164	10	AC146698	AC146698 Mus muscu	C 346	27	96.4	152044	2	AC147165	Pan trogl
C 274	27	96.4	122928	8	AC002387	AC002387 Arabidops	C 347	27	96.4	152044	2	AC147165	Pan trogl
C 275	27	96.4	124131	2	AC016335	AC016335 Homo sapi	C 348	27	96.4	152527	8	AC146716	Oryza sat
C 276	27	96.4	125491	1	AF550679	AF550679 Escherich	C 349	27	96.4	152668	2	AC119718	Rattus no
C 277	27	96.4	125656	8	AC146940	AC146940 Medicago	C 350	27	96.4	152845	8	AP005312	Oryza sat
C 278	27	96.4	126036	9	AC027312	AC027312 Homo sapi	C 351	27	96.4	153070	9	AL355361	Human DNA
C 279	27	96.4	127340	9	AC129907	AC129907 Homo sapi	C 352	27	96.4	153962	10	AL929149	Mouse DNA
C 280	27	96.4	128050	1	SAG766849	AL766849 Streptococ	C 353	27	96.4	154189	2	AC073180	Homo sapi
C 281	27	96.4	128293	10	AL928607	AL928607 Mouse DNA	C 354	27	96.4	154402	2	AC139634	Gallus ga
C 282	27	96.4	128330	9	HS111D6	AL301056 Human DNA	C 355	27	96.4	154572	2	AC120308	Oryza sat
C 283	27	96.4	128747	10	AL669846	AL669846 Mouse DNA	C 356	27	96.4	154590	9	AC105216	Homo sapi
C 284	27	96.4	129090	2	AC083961	AC083961 Homo sapi	C 357	27	96.4	155138	2	AC143259	Macaca mu
C 285	27	96.4	129585	8	AC138017	AC138017 Medicago	C 358	27	96.4	155313	2	AC068315	Homo sapi
C 286	27	96.4	130114	2	AF127414	AF127414 Homo sapi	C 359	27	96.4	157007	2	AC066599	Homo sapi
C 287	27	96.4	131095	2	AL590071	AL590071 Homo sapi	C 360	27	96.4	157015	2	AP001365	Homo sapi
C 288	27	96.4	131993	9	AC137561	AC137561 Homo sapi	C 361	27	96.4	157260	10	AL731857	Mouse DNA
C 289	27	96.4	133221	9	AC104841	AC104841 Homo sapi	C 362	27	96.4	157564	2	AC025996	Homo sapi
C 290	27	96.4	133565	9	AL357274	AL357274 Human DNA	C 363	27	96.4	157845	5	BX005079	Zebrafish
C 291	27	96.4	133690	9	AL158074	AL158074 Human DNA	C 364	27	96.4	158119	9	AC079098	Homo sapi
C 292	27	96.4	134097	8	AC135422	AC135422 Oryza sat	C 365	27	96.4	158458	2	AC091495	Homo sapi
C 293	27	96.4	134365	9	AC012624	AC012624 Homo sapi	C 366	27	96.4	158649	2	CR391916	Danio rer
C 294	27	96.4	136495	9	AC005521	AC005521 Homo sapi	C 367	27	96.4	158758	2	AC120375	Mus muscu
C 295	27	96.4	136725	9	AC026369	AC026369 Homo sapi	C 368	27	96.4	158849	9	HS435C23	Human DNA s
C 296	27	96.4	137327	8	AC135206	AC135206 Oryza sat	C 369	27	96.4	159177	9	AC147671	Pan trogl

C 370	27	96.4	159914	9	AC027553	AC027553 Homo sapi	C 443	27	96.4	174869	2	AP001840	AP001840 Homo sapi
C 371	27	96.4	160602	9	AC147661	AC147661 Pan trogl	C 444	27	96.4	174872	2	AC091571	AC091571 Homo sapi
C 372	27	96.4	160605	10	AC136101	AC136101 Rattus no	C 445	27	96.4	174874	2	AC026259	AC026259 Homo sapi
C 373	27	96.4	160776	2	AC137960	AC137960 Mus muscu	C 446	27	96.4	174892	2	AL135795	AL135795 Homo sapi
C 374	27	96.4	161326	9	AC004067	AC004067 Homo sapi	C 447	27	96.4	175674	2	AC053465	AC053465 Homo sapi
C 375	27	96.4	161479	8	AC090482	AC090482 Genomic s	C 448	27	96.4	176590	10	AC095688	AC095688 Rattus no
C 376	27	96.4	161518	2	AC012389	AC012389 Drosophil	C 449	27	96.4	176653	2	AP001857	AP001857 Homo sapi
C 377	27	96.4	161600	10	BX004866	BX004866 Mouse DNA	C 450	27	96.4	176714	9	AC005618	AC005618 Homo sapi
C 378	27	96.4	161771	9	CNS01DRX	AL121579 Human chr	C 451	27	96.4	176734	9	HSB395L14	AL078621 Human DNA
C 379	27	96.4	162062	10	AL691496	AL691496 Mouse DNA	C 452	27	96.4	177202	2	AC118703	AC118703 Mus muscu
C 380	27	96.4	162135	2	AC139182	AC139182 Mus muscu	C 453	27	96.4	177363	9	AC091051	AC091051 Homo sapi
C 381	27	96.4	162158	2	BX936466	BX936466 Danio rer	C 454	27	96.4	177431	9	AP001458	AP001458 Homo sapi
C 382	27	96.4	162190	2	AC020592	AC020592 Homo sapi	C 455	27	96.4	177481	9	AC149097	AC149097 Pan trogl
C 383	27	96.4	162556	10	AL672228	AL672228 Mouse DNA	C 456	27	96.4	177563	2	AC090336	AC090336 Homo sapi
C 384	27	96.4	162798	2	BX321900	BX321900 Danio rer	C 457	27	96.4	177647	9	AC091969	AC091969 Homo sapi
C 385	27	96.4	162897	2	AC110880	AC110880 Mus muscu	C 458	27	96.4	177669	9	CNS05TE7	AL358336 Human chr
C 386	27	96.4	163195	5	AL844172	AL844172 Zebrafish	C 459	27	96.4	177788	9	AC007342	AC007342 Homo sapi
C 387	27	96.4	163253	9	AL121877	AL121877 Human DNA	C 460	27	96.4	177790	10	AC140459	AC140459 Mus muscu
C 388	27	96.4	163409	2	AC123745	AC123745 Mus muscu	C 461	27	96.4	177792	10	AC140459	AC140459 Mus muscu
C 389	27	96.4	163489	2	AC025078	AC025078 Homo sapi	C 462	27	96.4	177792	10	AC119486	AC119486 Rattus no
C 390	27	96.4	163758	2	AC141345	AC141345 Rattus no	C 463	27	96.4	177885	2	AC128258	AC128258 Rattus no
C 391	27	96.4	163774	10	AC097355	AC097355 Mus muscu	C 464	27	96.4	177990	2	AC090306	AC090306 Homo sapi
C 392	27	96.4	163869	9	AL356967	AL356967 Human DNA	C 465	27	96.4	178604	2	AC149624	AC149624 Papio anu
C 393	27	96.4	164747	2	AC134191	AC134191 Rattus no	C 466	27	96.4	178728	2	AC145971	AC145971 Pan trogl
C 394	27	96.4	164798	9	AC092598	AC092598 Homo sapi	C 467	27	96.4	178830	8	AP003286	AP003286 Oryza sat
C 395	27	96.4	164936	2	AC141529	AC141529 Rattus no	C 468	27	96.4	178902	2	AL513525	AL513525 Homo sapi
C 396	27	96.4	165016	2	AC021530	AC021530 Homo sapi	C 469	27	96.4	179098	9	AP001521	AP001521 Homo sapi
C 397	27	96.4	165066	8	AE017078	AE017078 Oryza sat	C 470	27	96.4	179187	9	AC073583	AC073583 Homo sapi
C 398	27	96.4	165232	10	AL831731	AL831731 Mouse DNA	C 471	27	96.4	179216	10	AC104832	AC104832 Mus muscu
C 399	27	96.4	165496	9	AC104461	AC104461 Homo sapi	C 472	27	96.4	179385	9	AC093302	AC093302 Homo sapi
C 400	27	96.4	165671	2	CR450795	CR450795 Danio rer	C 473	27	96.4	179447	10	AL732454	AL732454 Mouse DNA
C 401	27	96.4	165673	2	AL354765	AL354765 Homo sapi	C 474	27	96.4	179728	9	AC091320	AC091320 Homo sapi
C 402	27	96.4	165690	2	AL162452	AL162452 Homo sapi	C 475	27	96.4	179767	9	AC069223	AC069223 Homo sapi
C 403	27	96.4	165789	5	AL953861	AL953861 Zebrafish	C 476	27	96.4	179930	5	EX537286	EX537286 Zebrafish
C 404	27	96.4	166537	2	AC139896	AC139896 Rattus no	C 477	27	96.4	180566	9	AP002409	AP002409 Homo sapi
C 405	27	96.4	166743	2	AC1410242	AC1410242 Mus muscu	C 478	27	96.4	180638	9	AC068763	AC068763 Homo sapi
C 406	27	96.4	166862	9	CNS01DTS	AL132857 Human chr	C 479	27	96.4	180674	2	AC148565	AC148565 Papio anu
C 407	27	96.4	166886	8	AP003248	AP003248 Oryza sat	C 480	27	96.4	180726	10	AC125116	AC125116 Mus muscu
C 408	27	96.4	166904	9	AL627309	AL627309 Human DNA	C 481	27	96.4	180797	9	AC147130	AC147130 Pan trogl
C 409	27	96.4	166965	10	AL831740	AL831740 Mouse DNA	C 482	27	96.4	180968	2	AC078996	AC078996 Mus muscu
C 410	27	96.4	167302	2	AC113980	AC113980 Mus muscu	C 483	27	96.4	181112	2	AL360012	AL360012 Homo sapi
C 411	27	96.4	167597	10	AC123551	AC123551 Mus muscu	C 484	27	96.4	181222	2	AP001530	AP001530 Homo sapi
C 412	27	96.4	167738	9	AC147056	AC147056 Pan trogl	C 485	27	96.4	181240	2	AC140725	AC140725 Homo sapi
C 413	27	96.4	168716	9	AC021299	AC021299 Homo sapi	C 486	27	96.4	181352	2	AC101973	AC101973 Mus muscu
C 414	27	96.4	168803	9	AC142350	AC142350 Pan trogl	C 487	27	96.4	181925	9	AC107877	AC107877 Homo sapi
C 415	27	96.4	168850	9	AP005364	AP005364 Homo sapi	C 488	27	96.4	181952	9	AC017053	AC017053 Homo sapi
C 416	27	96.4	168935	10	AC130715	AC130715 Mus muscu	C 489	27	96.4	182010	2	AC016767	AC016767 Homo sapi
C 417	27	96.4	169442	10	AC124545	AC124545 Homo sapi	C 490	27	96.4	182262	2	AC109195	AC109195 Mus muscu
C 418	27	96.4	169578	9	AC018522	AC018522 Homo sapi	C 491	27	96.4	182271	9	AC147667	AC147667 Pan trogl
C 419	27	96.4	169713	2	AC016733	AC016733 Homo sapi	C 492	27	96.4	182452	9	AL161731	AL161731 Human DNA
C 420	27	96.4	169931	9	AC005822	AC005822 Homo sapi	C 493	27	96.4	182529	9	AC092813	AC092813 Homo sapi
C 421	27	96.4	170533	9	AC148868	AC148868 Pan trogl	C 494	27	96.4	183155	2	AC128618	AC128618 Rattus no
C 422	27	96.4	170885	10	AC126933	AC126933 Mus muscu	C 495	27	96.4	183441	10	AC079956	AC079956 Mus muscu
C 423	27	96.4	171179	2	AC034140	AC034140 Homo sapi	C 496	27	96.4	183513	2	AC104753	AC104753 Mus muscu
C 424	27	96.4	171310	2	AC127934	AC127934 Rattus no	C 497	27	96.4	183527	9	AC146168	AC146168 Pan trogl
C 425	27	96.4	171555	2	AC087537	AC087537 Homo sapi	C 498	27	96.4	183571	2	AC102665	AC102665 Mus muscu
C 426	27	96.4	171641	9	AC018512	AC018512 Homo sapi	C 499	27	96.4	184012	9	AC017080	AC017080 Homo sapi
C 427	27	96.4	172040	2	AC142192	AC142192 Rattus no	C 500	27	96.4	184342	2	AC122209	AC122209 Mus muscu
C 428	27	96.4	172421	2	AC013279	AC013279 Homo sapi	C 501	27	96.4	184428	2	AC119645	AC119645 Rattus no
C 429	27	96.4	173073	9	AC099522	AC099522 Homo sapi	C 502	27	96.4	184557	10	AC036111	AC036111 Homo sapi
C 430	27	96.4	173238	2	AC149017	AC149017 Papio anu	C 503	27	96.4	184570	9	AC036111	AC036111 Homo sapi
C 431	27	96.4	173602	9	AC090226	AC090226 Homo sapi	C 504	27	96.4	184740	2	AC107766	AC107766 Mus muscu
C 432	27	96.4	173602	10	AC132222	AC132222 Mus muscu	C 505	27	96.4	184855	10	AC124713	AC124713 Mus muscu
C 433	27	96.4	173772	2	AC087862	AC087862 Homo sapi	C 506	27	96.4	184866	2	AC148420	AC148420 Melagris
C 434	27	96.4	173949	10	AC132100	AC132100 Mus muscu	C 507	27	96.4	185067	9	AC022382	AC022382 Homo sapi
C 435	27	96.4	173967	9	AC147338	AC147338 Pan trogl	C 508	27	96.4	185377	2	AC110036	AC110036 Mus muscu
C 436	27	96.4	174023	9	AC080089	AC080089 Homo sapi	C 509	27	96.4	185800	2	AC135224	AC135224 Gallus ga
C 437	27	96.4	174285	2	AC150279	AC150279 Pan trogl	C 510	27	96.4	185995	2	AC087306	AC087306 Homo sapi
C 438	27	96.4	174285	2	AC150279	AC150279 Pan trogl	C 511	27	96.4	186614	1	AP003602	AP003602 Nostoc ep
C 439	27	96.4	174361	9	AL158044	AL158044 Human DNA	C 512	27	96.4	186984	9	AC109496	AC109496 Homo sapi
C 440	27	96.4	174478	8	AP005467	AP005467 Oryza sat	C 513	27	96.4	187140	9	AC146186	AC146186 Pan trogl
C 441	27	96.4	174535	9	AC090465	AC090465 Homo sapi	C 514	27	96.4	187233	10	AC107661	AC107661 Mus muscu
C 442	27	96.4	174582	2	AC116929	AC116929 Rattus no	C 515	27	96.4	187739	10	AC109503	AC109503 Mus muscu

516	27	96.4	187741	2	AC087681	AC087681 Homo sapi	c 589	27	96.4	204445	2	AC126512	AC126512 Rattus no
517	27	96.4	188850	9	AC146578	AC146578 Pan trogl	590	27	96.4	205779	9	BS000597	BS000597 Pan trogl
518	27	96.4	188853	10	AC102105	AC102105 Mus muscu	591	27	96.4	206324	10	AC122798	AC122798 Mus muscu
519	27	96.4	189042	9	AC146245	AC146245 Pan trogl	c 592	27	96.4	206832	2	AC107729	AC107729 Mus muscu
520	27	96.4	189058	2	CR388373	CR388373 Danio rer	c 593	27	96.4	207685	2	BX572631	BX572631 Danio rer
521	27	96.4	189098	9	AL512430	AL512430 Human DNA	c 594	27	96.4	207736	2	AC131964	AC131964 Rattus no
522	27	96.4	189158	9	AC147666	AC147666 Pan trogl	c 595	27	96.4	207768	10	AC099735	AC099735 Mus muscu
523	27	96.4	189239	10	AC105969	AC105969 Mus muscu	596	27	96.4	207875	2	AC130657	AC130657 Mus muscu
524	27	96.4	189430	2	AC011610	AC011610 Homo sapi	c 597	27	96.4	207922	2	AC073790	AC073790 Mus muscu
525	27	96.4	189467	2	AC109218	AC109218 Mus muscu	598	27	96.4	208596	2	AC127589	AC127589 Mus muscu
526	27	96.4	189495	10	AL607083	AL607083 Mouse DNA	599	27	96.4	208615	2	AC079180	AC079180 Mus muscu
527	27	96.4	189616	2	AC121093	AC121093 Mus muscu	c 600	27	96.4	209038	2	AC113280	AC113280 Mus muscu
528	27	96.4	189679	2	AC022337	AC022337 Homo sapi	601	27	96.4	209172	2	AC101950	AC101950 Mus muscu
529	27	96.4	189946	2	AC120255	AC120255 Rattus no	602	27	96.4	209705	2	AC007931	AC007931 Homo sapi
530	27	96.4	190048	10	AL591126	AL591126 Mouse DNA	c 603	27	96.4	209762	10	AC114822	AC114822 Mus muscu
531	27	96.4	190145	2	AC147334	AC147334 Pan trogl	604	27	96.4	209780	2	AC102649	AC102649 Mus muscu
532	27	96.4	190374	2	AC132647	AC132647 Rattus no	c 605	27	96.4	210384	2	CR556714	CR556714 Danio rer
533	27	96.4	190916	10	AC122470	AC122470 Mus muscu	606	27	96.4	211018	3	CNS078GC	AL590445 chromosom
534	27	96.4	191310	10	AC079365	AC079365 Mus muscu	c 607	27	96.4	211034	10	AL731709	AL731709 Mouse DNA
535	27	96.4	191331	10	AC131325	AC131325 Mus muscu	c 608	27	96.4	211250	2	AC090702	AC090702 Homo sapi
536	27	96.4	191957	10	AL606922	AL606922 Mouse DNA	609	27	96.4	211796	10	AC131728	AC131728 Mus muscu
537	27	96.4	191957	10	AL606922	AL606922 Mouse DNA	610	27	96.4	211910	10	AC116573	AC116573 Mus muscu
538	27	96.4	192177	2	AC147674	AC147674 Pan trogl	611	27	96.4	212299	2	AC110628	AC110628 Rattus no
539	27	96.4	192207	10	AC105950	AC105950 Mus muscu	c 612	27	96.4	212365	9	AC147687	AC147687 Pan trogl
540	27	96.4	192253	2	AC020615	AC020615 Homo sapi	c 613	27	96.4	212472	10	AL731838	AL731838 Mouse DNA
541	27	96.4	192484	10	AC134906	AC134906 Mus muscu	c 614	27	96.4	213015	2	AC128800	AC128800 Rattus no
542	27	96.4	192556	2	AC149121	AC149121 Papio anu	c 615	27	96.4	213032	10	AC117679	AC117679 Mus muscu
543	27	96.4	192748	9	AC0011089	AC0011089 Homo sapi	616	27	96.4	213110	2	AC087324	AC087324 Mus muscu
544	27	96.4	192966	9	AL353813	AL353813 Human DNA	617	27	96.4	213244	10	AC113016	AC113016 Mus muscu
545	27	96.4	193586	2	AC139447	AC139447 Rattus no	c 618	27	96.4	213540	2	AC112652	AC112652 Homo sapi
546	27	96.4	193944	4	AC147679	AC147679 Canis Fam	c 619	27	96.4	213746	9	AC146190	AC146190 Pan trogl
547	27	96.4	194181	9	AC124148	AC124148 Pan trogl	c 620	27	96.4	213908	2	AC113909	AC113909 Rattus no
548	27	96.4	194313	5	BX323807	BX323807 Zebrafish	621	27	96.4	213918	10	AL732541	AL732541 Mouse DNA
549	27	96.4	194441	2	BX664703	BX664703 Danio rer	c 622	27	96.4	214118	9	BS000047	BS000047 Pan trogl
550	27	96.4	194563	2	AL391687	AL391687 Homo sapi	623	27	96.4	214489	10	AC145554	AC145554 Mus muscu
551	27	96.4	194696	2	AC148941	AC148941 Pan trogl	c 624	27	96.4	215067	2	AC114214	AC114214 Rattus no
552	27	96.4	194707	2	AC128923	AC128923 Rattus no	c 625	27	96.4	215273	5	BX248502	BX248502 Zebrafish
553	27	96.4	194753	2	AC128032	AC128032 Rattus no	c 626	27	96.4	215713	2	AC102762	AC102762 Rattus no
554	27	96.4	195273	10	AC147474	AC147474 Mus muscu	c 627	27	96.4	216502	2	AC109509	AC109509 Mus muscu
555	27	96.4	195433	9	AC026778	AC026778 Homo sapi	628	27	96.4	218117	2	AC098140	AC098140 Rattus no
556	27	96.4	195978	10	AC111099	AC111099 Mus muscu	629	27	96.4	218583	2	AC121165	AC121165 Rattus no
557	27	96.4	195937	10	BX119986	BX119986 Mouse DNA	c 630	27	96.4	218678	2	AC123117	AC123117 Rattus no
558	27	96.4	197186	2	AC128937	AC128937 Rattus no	c 631	27	96.4	218778	2	AC137406	AC137406 Rattus no
559	27	96.4	197988	2	AC112849	AC112849 Rattus no	c 632	27	96.4	218859	2	AC022912	AC022912 Homo sapi
560	27	96.4	198019	10	AC108944	AC108944 Mus muscu	633	27	96.4	218921	2	AC097912	AC097912 Rattus no
561	27	96.4	198087	9	AC008074	AC008074 Homo sapi	c 634	27	96.4	219258	9	AC022121	AC022121 Homo sapi
562	27	96.4	198167	9	AC093528	AC093528 Homo sapi	c 635	27	96.4	219574	9	AC016759	AC016759 Homo sapi
563	27	96.4	198365	5	BX323389	BX323389 Zebrafish	636	27	96.4	219809	2	AF336379	AF336379 Mus muscu
564	27	96.4	198378	2	AC103657	AC103657 Mus muscu	c 637	27	96.4	219991	9	BS000547	BS000547 Pan trogl
565	27	96.4	198582	9	AC005291	AC005291 Homo sapi	c 638	27	96.4	221534	10	AL929137	AL929137 Mouse DNA
566	27	96.4	198700	9	AC104317	AC104317 Homo sapi	c 639	27	96.4	222444	2	AC097876	AC097876 Rattus no
567	27	96.4	198832	2	AC101759	AC101759 Mus muscu	640	27	96.4	222490	9	AL353801	AL353801 Human DNA
568	27	96.4	198899	9	AC090192	AC090192 Homo sapi	c 641	27	96.4	222512	9	AC006335	AC006335 Homo sapi
569	27	96.4	199415	10	AC084053	AC084053 Mus Muscu	c 642	27	96.4	222606	2	AC125407	AC125407 Mus muscu
570	27	96.4	199457	8	AP003407	AP003407 Oryza sat	c 643	27	96.4	222920	2	AC098425	AC098425 Rattus no
571	27	96.4	199912	9	AC008430	AC008430 Homo sapi	644	27	96.4	223157	2	AC148972	AC148972 Mus muscu
572	27	96.4	200064	2	AC150681	AC150681 Mus muscu	c 645	27	96.4	223171	2	AC115753	AC115753 Mus muscu
573	27	96.4	200244	2	AC109944	AC109944 Rattus no	646	27	96.4	224108	10	AC122930	AC122930 Mus muscu
574	27	96.4	200272	2	AC134059	AC134059 Rattus no	647	27	96.4	224290	2	AC098199	AC098199 Rattus no
575	27	96.4	200536	9	AL356093	AL356093 Human DNA	c 648	27	96.4	225171	2	AC118251	AC118251 Mus muscu
576	27	96.4	200572	2	AC141398	AC141398 Rattus no	649	27	96.4	225338	2	AC106924	AC106924 Rattus no
577	27	96.4	200577	1	BSUB0015	Z99118 Bacillus su	c 650	27	96.4	225419	1	ECOUW76	U00039 E. coli chr
578	27	96.4	201416	2	AC130437	AC130437 Homo sapi	651	27	96.4	225422	2	AC097234	AC097234 Rattus no
579	27	96.4	201747	2	AC119135	AC119135 Rattus no	652	27	96.4	225555	2	AC144789	AC144789 Gallus ga
580	27	96.4	201783	10	AC120148	AC120148 Mus muscu	653	27	96.4	225803	2	AC123641	AC123641 Mus muscu
581	27	96.4	201894	2	AC009504	AC009504 Homo sapi	c 654	27	96.4	225868	10	AC098743	AC098743 Mus muscu
582	27	96.4	202402	2	AC012517	AC012517 Homo sapi	c 655	27	96.4	225989	2	AC135294	AC135294 Rattus no
583	27	96.4	202505	2	AC090232	AC090232 Homo sapi	656	27	96.4	226142	2	AC126422	AC126422 Mus muscu
584	27	96.4	202611	9	AL731809	AL731809 Human DNA	c 657	27	96.4	226667	2	AC107430	AC107430 Rattus no
585	27	96.4	202611	9	AL731809	AL731809 Human DNA	658	27	96.4	227155	10	AL732593	AL732593 Mouse DNA
586	27	96.4	203555	2	AC149049	AC149049 Papio anu	659	27	96.4	227202	2	AC128501	AC128501 Rattus no
587	27	96.4	203841	2	AC141954	AC141954 Rattus no	660	27	96.4	227564	2	CR533429	CR533429 Danio rer
588	27	96.4	203856	2	AC111387	AC111387 Rattus no	661	27	96.4	227866	2	AC126744	AC126744 Rattus no

c 662	27	96.4	227977	10	AC118561	AC118561 Mus muscu	c 735	27	96.4	250318	2	AC109983	AC109983 Rattus no
663	27	96.4	229118	2	AC129880	AC129880 Rattus no	c 736	27	96.4	250839	2	AC115372	AC115372 Rattus no
664	27	96.4	229287	2	AC036172	AC036172 Homo sapi	737	27	96.4	250840	2	AC127193	AC127193 Rattus no
c 665	27	96.4	229315	2	AC111736	AC111736 Rattus no	738	27	96.4	251132	2	AC098162	AC098162 Rattus no
666	27	96.4	230130	10	AC112633	AC112633 Rattus no	c 739	27	96.4	251448	3	AE014819	AE014819 Plasmodiu
c 667	27	96.4	230561	10	AC131172	AC131172 Rattus no	740	27	96.4	251945	2	AC103124	AC103124 Rattus no
668	27	96.4	231224	2	AC131466	AC131466 Rattus no	c 741	27	96.4	252215	2	AC123125	AC123125 Rattus no
c 669	27	96.4	231660	2	AC108236	AC108236 Rattus no	c 742	27	96.4	252448	2	AC094385	AC094385 Rattus no
670	27	96.4	231741	2	AC110855	AC110855 Rattus no	c 743	27	96.4	252612	2	AC111263	AC111263 Rattus no
671	27	96.4	231803	2	AC104908	AC104908 Mus muscu	744	27	96.4	253140	2	AC094055	AC094055 Rattus no
672	27	96.4	231916	2	AC137175	AC137175 Rattus no	c 745	27	96.4	253361	2	AC112312	AC112312 Rattus no
c 673	27	96.4	231938	2	AC079025	AC079025 Mus muscu	746	27	96.4	253719	2	AC105548	AC105548 Rattus no
674	27	96.4	232264	2	AC130474	AC130474 Mus muscu	747	27	96.4	253744	10	AC107368	AC107368 Mus muscu
c 675	27	96.4	232501	10	AC115410	AC115410 Rattus no	748	27	96.4	253960	2	AC094858	AC094858 Rattus no
676	27	96.4	232631	2	AC136122	AC136122 Rattus no	749	27	96.4	254400	2	AC106171	AC106171 Rattus no
677	27	96.4	233572	2	AC094202	AC094202 Rattus no	c 750	27	96.4	254884	2	AC098430	AC098430 Rattus no
678	27	96.4	233659	2	CR545476	CR545476 Danio rer	c 751	27	96.4	255608	2	AC109405	AC109405 Rattus no
c 679	27	96.4	234483	2	AC098407	AC098407 Rattus no	c 752	27	96.4	256086	2	AC097863	AC097863 Rattus no
c 680	27	96.4	234795	10	AC099934	AC099934 Mus muscu	753	27	96.4	256200	2	AC132693	AC132693 Rattus no
c 681	27	96.4	235317	2	AC121371	AC121371 Rattus no	754	27	96.4	256354	10	AC120922	AC120922 Rattus no
c 682	27	96.4	235445	2	AC116743	AC116743 Mus muscu	755	27	96.4	258510	2	AC130514	AC130514 Rattus no
c 683	27	96.4	235993	2	AC108558	AC108558 Rattus no	756	27	96.4	258581	2	AC117354	AC117354 Rattus no
684	27	96.4	236094	2	AC126149	AC126149 Rattus no	757	27	96.4	259203	2	AC097682	AC097682 Rattus no
685	27	96.4	236533	2	AC103433	AC103433 Rattus no	c 758	27	96.4	259900	2	AC133278	AC133278 Rattus no
686	27	96.4	236624	2	AC103434	AC103434 Rattus no	c 759	27	96.4	260888	2	AC106440	AC106440 Rattus no
c 687	27	96.4	236661	10	AC116744	AC116744 Mus muscu	760	27	96.4	261688	2	AC111685	AC111685 Rattus no
c 688	27	96.4	236979	10	AL591478	AL591478 Mouse DNA	761	27	96.4	261701	2	AC111897	AC111897 Rattus no
c 689	27	96.4	237247	10	AC095563	AC095563 Rattus no	c 762	27	96.4	261847	2	AC130670	AC130670 Mus muscu
690	27	96.4	237271	2	AC106162	AC106162 Rattus no	763	27	96.4	261990	2	AC111513	AC111513 Rattus no
691	27	96.4	237468	2	AC108532	AC108532 Rattus no	764	27	96.4	262034	2	AC112555	AC112555 Rattus no
692	27	96.4	237517	2	AC107160	AC107160 Rattus no	765	27	96.4	263827	2	AC094579	AC094579 Rattus no
693	27	96.4	237617	2	AC148938	AC148938 Pan trogl	c 766	27	96.4	264095	9	AC008500	AC008500 Homo sapi
694	27	96.4	237678	10	AC055818	AC055818 Mus muscu	c 767	27	96.4	264608	2	AC111627	AC111627 Rattus no
c 695	27	96.4	238042	2	AC132991	AC132991 Rattus no	c 768	27	96.4	265103	2	AC129636	AC129636 Rattus no
c 696	27	96.4	238397	2	AC137421	AC137421 Rattus no	769	27	96.4	266331	2	AC142263	AC142263 Mus muscu
c 697	27	96.4	238516	10	AC118928	AC118928 Mus muscu	770	27	96.4	268150	1	AC119310	AC119310 Rattus no
c 698	27	96.4	238816	2	AC108616	AC108616 Rattus no	c 771	27	96.4	268356	1	TACID5	AL445067 Thermopla
c 699	27	96.4	239292	2	AC108329	AC108329 Rattus no	c 772	27	96.4	270139	2	AC128450	AC128450 Rattus no
c 700	27	96.4	239348	2	AC129136	AC129136 Rattus no	773	27	96.4	273331	2	AC110676	AC110676 Rattus no
c 701	27	96.4	239367	2	AC103462	AC103462 Rattus no	c 774	27	96.4	274762	1	EX664015	EX664015 Serratia
c 702	27	96.4	239416	2	AC130914	AC130914 Rattus no	c 775	27	96.4	275710	2	AC135276	AC135276 Rattus no
c 703	27	96.4	239684	2	AC096619	AC096619 Rattus no	c 776	27	96.4	276181	2	AC092382	AC092382 Homo sapi
c 704	27	96.4	240103	2	AC111794	AC111794 Rattus no	c 777	27	96.4	276450	2	AC111593	AC111593 Rattus no
705	27	96.4	240389	2	AC107489	AC107489 Rattus no	c 778	27	96.4	276496	2	AC091355	AC091355 Rattus no
c 706	27	96.4	240482	2	AC122975	AC122975 Rattus no	c 779	27	96.4	280900	1	AP002566	AP002566 Escherich
707	27	96.4	240729	2	AC111961	AC111961 Rattus no	c 780	27	96.4	280900	2	AC132748	AC132748 Rattus no
708	27	96.4	241048	2	AC129824	AC129824 Rattus no	781	27	96.4	281324	2	AC103308	AC103308 Rattus no
709	27	96.4	241213	2	AC139952	AC139952 Rattus no	c 782	27	96.4	281752	2	EX088646	EX088646 Danio rer
c 710	27	96.4	241571	2	AC108281	AC108281 Rattus no	783	27	96.4	281917	2	AC102191	AC102191 Mus muscu
c 711	27	96.4	241807	2	AC107493	AC107493 Rattus no	784	27	96.4	283789	2	AC110840	AC110840 Rattus no
c 712	27	96.4	242037	2	AC114200	AC114200 Rattus no	785	27	96.4	287044	2	AC149054	AC149054 Mus muscu
c 713	27	96.4	242093	2	AC123090	AC123090 Rattus no	c 786	27	96.4	287130	1	CNSPAX04	AJ248286 Pyrococcu
714	27	96.4	242123	2	AC094773	AC094773 Rattus no	c 787	27	96.4	287765	1	AE017276	AE017276 Bacillus
715	27	96.4	242618	2	AC103444	AC103444 Rattus no	788	27	96.4	288697	2	AC133236	AC133236 Rattus no
716	27	96.4	242738	2	AC114179	AC114179 Rattus no	789	27	96.4	289251	2	AC114658	AC114658 Mus muscu
717	27	96.4	243963	2	AC097147	AC097147 Rattus no	c 790	27	96.4	289758	2	AC105607	AC105607 Rattus no
718	27	96.4	244017	2	AC137168	AC137168 Rattus no	c 791	27	96.4	291030	1	AE017036	AE017036 Bacillus
c 719	27	96.4	244170	2	AC130152	AC130152 Rattus no	792	27	96.4	293558	1	AE016991	AE016991 Shigella
c 720	27	96.4	245067	2	AC087136	AC087136 Mus muscu	793	27	96.4	295920	2	AC108611	AC108611 Rattus no
721	27	96.4	245095	2	AC111774	AC111774 Rattus no	794	27	96.4	295970	2	AC128907	AC128907 Rattus no
722	27	96.4	245297	2	AC126625	AC126625 Rattus no	795	27	96.4	296650	1	AP005082	AP005082 Vibrio pa
c 723	27	96.4	246612	2	AC008579	AC008579 Homo sapi	796	27	96.4	300029	8	AE017094	AE017094 Oryza sat
c 724	27	96.4	246642	2	AC123079	AC123079 Rattus no	797	27	96.4	300242	1	AE016790	AE016790 Pseudomon
c 725	27	96.4	246829	2	AC121465	AC121465 Rattus no	798	27	96.4	300337	1	AE016927	AE016927 Bacteroid
c 726	27	96.4	246883	2	AC126110	AC126110 Rattus no	c 799	27	96.4	301479	9	AC136590	AC136590 Homo sapi
727	27	96.4	247100	2	AC100967	AC100967 Mus muscu	c 800	27	96.4	301660	1	AC101678	AC101678 Escherich
c 728	27	96.4	247810	2	AC094400	AC094400 Rattus no	801	27	96.4	302050	1	EX321856	EX321856 Nitrosomo
c 729	27	96.4	247871	2	AC106902	AC106902 Rattus no	c 802	27	96.4	302481	1	AE017010	AE017010 Bacillus
730	27	96.4	247996	2	AC095789	AC095789 Rattus no	c 803	27	96.4	303414	1	AE015938	AE015938 Clostridi
c 731	27	96.4	248333	2	AC097969	AC097969 Rattus no	c 804	27	96.4	303450	1	AP005085	AP005085 Vibrio pa
732	27	96.4	249003	3	TBBCHRIA4	AL929607 Trypanoso	c 805	27	96.4	304050	1	AP004829	AP004829 Staphyloc
733	27	96.4	249240	2	AC098528	AC098528 Rattus no	c 806	27	96.4	305541	1	AE017290	AE017290 Leprosplr
c 734	27	96.4	249738	2	AC111644	AC111644 Rattus no	c 807	27	96.4	306214	2	AC123205	AC123205 Rattus no

C 808	27	96.4	307717	2	AC108630	Rattus no	881	26	92.9	273	6	CQ080959	Sequence
C 809	27	96.4	307750	1	AP003136	Staphyloc	882	26	92.9	273	6	CQ115437	Sequence
C 810	27	96.4	309838	2	AC118433	Rattus no	883	26	92.9	273	6	CQ154254	Sequence
C 811	27	96.4	315079	1	MPULM03	AL445565	884	26	92.9	273	6	CQ186830	Sequence
C 812	27	96.4	318312	2	AC097988	Myco plas	885	26	92.9	273	6	CQ237496	Sequence
C 813	27	96.4	328451	2	AC128741	Rattus no	886	26	92.9	273	6	CQ275120	Sequence
C 814	27	96.4	328822	2	AC125301	Rattus no	887	26	92.9	273	6	CQ343488	Sequence
C 815	27	96.4	329972	2	AC094626	Rattus no	888	26	92.9	275	11	BV073083	Sequence
C 816	27	96.4	342350	1	AP003365	Staphyloc	889	26	92.9	276	6	CQ12161	Sequence
C 817	27	96.4	343587	2	AC113471	Mus muscu	890	26	92.9	283	6	CQ724820	Sequence
C 818	27	96.4	346416	1	BX842648	Bdellovib	891	26	92.9	286	11	BV026481	Sequence
C 819	27	96.4	347400	1	AP003591	Nostoc sp	892	26	92.9	305	6	AX906181	Sequence
C 820	27	96.4	347800	1	AP000060	Aeropyrum	893	26	92.9	305	6	BD041714	Sequence
C 821	27	96.4	347800	1	AP000060	Aeropyrum	894	26	92.9	305	6	BD057527	Sequence
C 822	27	96.4	348077	1	AP000300	Mesorhizo	895	26	92.9	319	11	HS309XB1	Sequence
C 823	27	96.4	348551	8	BX649606	Aspergill	896	26	92.9	324	6	AR347564	Sequence
C 824	27	96.4	348813	1	BX571865	Photorhab	897	26	92.9	335	8	AF305102	Sequence
C 825	27	96.4	349260	1	BX572595	Rhodospir	898	26	92.9	336	6	AX110472	Sequence
C 826	27	96.4	349876	1	BX640442	Bordetell	899	26	92.9	336	8	CNS019WN	Sequence
C 827	27	96.4	349980	6	CQ655070	Sequence	900	26	92.9	336	8	CNS01BI7	Sequence
C 828	27	96.4	349980	6	AX041920	Sequence	901	26	92.9	340	6	CQ735908	Sequence
C 829	27	96.4	349980	6	AX954530	Sequence	902	26	92.9	340	6	CQ736611	Sequence
C 830	26	92.9	39	6	I32153	Sequence 7	903	26	92.9	343	6	CQ081397	Sequence
C 831	26	92.9	55	14	MUM3TER	Mumps virus	904	26	92.9	343	6	CQ115951	Sequence
C 832	26	92.9	65	6	CQ558469	Sequence	905	26	92.9	343	6	CQ115951	Sequence
C 833	26	92.9	107	6	AX920546	Sequence	906	26	92.9	343	6	CQ154719	Sequence
C 834	26	92.9	107	6	BD056079	Sequence	907	26	92.9	343	6	CQ187234	Sequence
C 835	26	92.9	112	9	HG187C1F	Sequence	908	26	92.9	343	6	CQ237966	Sequence
C 836	26	92.9	147	10	MUSBC12	D49927 Mouse DNA,	909	26	92.9	343	6	CQ275595	Sequence
C 837	26	92.9	150	6	CQ054608	Sequence	910	26	92.9	344	6	CQ503352	Sequence
C 838	26	92.9	150	6	CQ073851	Sequence	911	26	92.9	345	11	CR377251	Sequence
C 839	26	92.9	150	6	CQ104740	Sequence	912	26	92.9	347	6	CQ682645	Sequence
C 840	26	92.9	150	6	CQ134444	Sequence	913	26	92.9	349	14	AY146962	Sequence
C 841	26	92.9	150	6	CQ178936	Sequence	914	26	92.9	360	6	AX184730	Sequence
C 842	26	92.9	150	6	CQ203281	Sequence	915	26	92.9	362	6	CQ734126	Sequence
C 843	26	92.9	150	6	CQ226646	Sequence	916	26	92.9	367	11	BV091814	Sequence
C 844	26	92.9	150	6	CQ264788	Sequence	917	26	92.9	375	11	BV051841	Sequence
C 845	26	92.9	150	6	CQ301877	Sequence	918	26	92.9	378	5	AF085755	Sequence
C 846	26	92.9	150	6	CQ339075	Sequence	919	26	92.9	381	11	BX537463	Sequence
C 847	26	92.9	159	6	I32155	Sequence 9	920	26	92.9	383	6	AX919798	Sequence
C 848	26	92.9	165	6	CQ079500	Sequence	921	26	92.9	383	6	BD055331	Sequence
C 849	26	92.9	165	6	CQ111449	Sequence	922	26	92.9	387	6	AR394150	Sequence
C 850	26	92.9	165	6	CQ150231	Sequence	923	26	92.9	388	14	AY146964	Sequence
C 851	26	92.9	165	6	CQ184681	Sequence	924	26	92.9	390	14	AF028331	Sequence
C 852	26	92.9	165	6	CQ233536	Sequence	925	26	92.9	394	6	CQ075003	Sequence
C 853	26	92.9	165	6	CQ271408	Sequence	926	26	92.9	394	6	CQ105935	Sequence
C 854	26	92.9	165	6	CQ308857	Sequence	927	26	92.9	394	6	CQ144625	Sequence
C 855	26	92.9	165	6	CQ345574	Sequence	928	26	92.9	394	6	CQ180102	Sequence
C 856	26	92.9	170	6	AX917563	Sequence	929	26	92.9	394	6	CQ204475	Sequence
C 857	26	92.9	170	6	BD053096	Sequence	930	26	92.9	394	6	CQ227825	Sequence
C 858	26	92.9	171	9	D16495815	Sequence	931	26	92.9	394	6	CQ340258	Sequence
C 859	26	92.9	198	14	AY146966	Sequence	932	26	92.9	396	14	AY146963	Sequence
C 860	26	92.9	201	11	BV198407	sgm19744	933	26	92.9	402	6	CQ503226	Sequence
C 861	26	92.9	218	11	BV007508	MASC_STS1	934	26	92.9	403	6	AX899953	Sequence
C 862	26	92.9	219	11	BV007950	MASC_STS1	935	26	92.9	403	6	BD035486	Sequence
C 863	26	92.9	219	11	BV007951	MASC_STS1	936	26	92.9	403	8	AY285895	Sequence
C 864	26	92.9	219	11	BV007952	MASC_STS1	937	26	92.9	404	6	AX912563	Sequence
C 865	26	92.9	220	11	BV007954	MASC_STS1	938	26	92.9	404	6	BD048096	Sequence
C 866	26	92.9	230	11	BV007867	MASC_STS1	939	26	92.9	408	6	CQ080893	Sequence
C 867	26	92.9	232	8	AF327807	Fusarium	940	26	92.9	408	6	CQ115364	Sequence
C 868	26	92.9	237	14	AY146968	Sequence	941	26	92.9	408	6	CQ186767	Sequence
C 869	26	92.9	239	6	CQ682981	Sequence	942	26	92.9	408	6	CQ237428	Sequence
C 870	26	92.9	239	6	AR417793	Sequence	943	26	92.9	408	6	CQ275053	Sequence
C 871	26	92.9	239	6	AX978487	Sequence	944	26	92.9	408	6	CQ349417	Sequence
C 872	26	92.9	239	6	BD113346	EST and e	945	26	92.9	428	8	SP1489510	Sequence
C 873	26	92.9	245	14	AY146967	Sequence	946	26	92.9	430	8	AF492580	Sequence
C 874	26	92.9	252	4	AF088963	Sequence	947	26	92.9	435	6	AX963133	Sequence
C 875	26	92.9	258	9	AB181941	Sequence	948	26	92.9	444	6	CQ080653	Sequence
C 876	26	92.9	262	6	AX901570	Sequence	949	26	92.9	444	6	CQ115100	Sequence
C 877	26	92.9	262	6	BD037103	Sequence	950	26	92.9	444	6	CQ153945	Sequence
C 878	26	92.9	263	6	AX887480	Sequence	951	26	92.9	444	6	CQ186530	Sequence
C 879	26	92.9	263	6	BD027090	Sequence	952	26	92.9	444	6	CQ237176	Sequence
C 880	26	92.9	270	6	AY23304	Sequence 17	953	26	92.9	444	6	CQ274794	Sequence

954 26 92.9 444 6 CQ311782 Sequence
955 26 92.9 444 6 CQ349137 Sequence
956 26 92.9 444 8 AF319759 Streptogl
957 26 92.9 445 6 AR230550 Sequence
958 26 92.9 447 8 AF319741 Pluchea d
959 26 92.9 455 6 AR420990 Sequence
960 26 92.9 455 6 AX981684 Sequence
961 26 92.9 455 6 BD116543 EST and e
962 26 92.9 456 6 CQ070300 Sequence
963 26 92.9 456 6 CQ098316 Sequence
964 26 92.9 456 6 CQ137209 Sequence
965 26 92.9 456 6 CQ174796 Sequence
966 26 92.9 456 6 CQ220542 Sequence
967 26 92.9 456 6 CQ258837 Sequence
968 26 92.9 456 6 CQ296354 Sequence
969 26 92.9 456 6 CQ332838 Sequence
970 26 92.9 463 6 CQ0482091 Sequence
971 26 92.9 466 6 CQ049570 Sequence
972 26 92.9 466 6 CQ064596 Sequence
973 26 92.9 466 6 CQ091543 Sequence
974 26 92.9 466 6 CQ130369 Sequence
975 26 92.9 466 6 CQ168989 Sequence
976 26 92.9 466 6 CQ198112 Sequence
977 26 92.9 466 6 CQ213573 Sequence
978 26 92.9 466 6 CQ252153 Sequence
979 26 92.9 466 6 CQ289301 Sequence
980 26 92.9 466 6 CQ326288 Sequence
981 26 92.9 466 8 AB083397 Phaneroch
982 26 92.9 474 6 CQ683699 Sequence
983 26 92.9 479 6 A42306 Sequence 19
984 26 92.9 479 11 G96965 S210P6227RC
985 26 92.9 480 6 AX507918 Sequence
986 26 92.9 480 8 AX098951 Arabidops
987 26 92.9 486 6 AX413779 Sequence
988 26 92.9 486 6 AX415697 Sequence
989 26 92.9 490 6 CQ740090 Sequence
990 26 92.9 495 6 AR263975 Sequence
991 26 92.9 497 11 BV017036 S208P6413
992 26 92.9 499 6 CQ399096 Sequence
993 26 92.9 499 6 CQ405374 Sequence
994 26 92.9 500 11 BV064607 S212P6568
995 26 92.9 502 14 AJ554762 Human imm
996 26 92.9 503 11 BV037059 S209P6501
997 26 92.9 505 6 AX187599 Sequence
998 26 92.9 505 11 CR379601 Arabidops
999 26 92.9 507 11 BX546264 Arabidops
1000 26 92.9 515 8 CNS01DH2 AL116814 Botrytis

ALIGNMENTS

RESULT 1
CQ713883 Sequence 58809 from Patent WO02070737. linear PAT 03-FEB-2004
LOCUS 200 bp DNA
DEFINITION CQ713883
ACCESSION CQ713883.1 GI:42274740
VERSION
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liew, C.C., Marshall, W.E. and Zhang, H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 58809 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source Location/Qualifiers
1..200
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

RESULT 3
I08256 Sequence 1 from Patent EP 0369567. linear PAT 02-DEC-1994
LOCUS 330 bp DNA
DEFINITION I08256
ACCESSION I08256
VERSION I08256.1 GI:589029
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 330)
AUTHORS Johnson, M.J
TITLE Chimeric antibodies directed against metal chelates
JOURNAL Patent: EP 0369567-A2 1 23-MAY-1990;
Location/Qualifiers
1..330
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:

Alignment Scores:
Pred. No.: 51.5 Length: 200
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x CQ713883 (1-200)
Qy 1 GlyTyr***ValGluGlu 6
Db 63 GGATATTCTGTGAAGAA 80

RESULT 2
AX310498 Sequence 3483 from Patent WO0190366. linear PAT 14-DEC-2001
LOCUS 318 bp DNA
DEFINITION AX310498
ACCESSION AX310498
VERSION AX310498.1 GI:17896479
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Leach, M.D. and Shinkets, R.A.
TITLE Human polynucleotides and polypeptides encoded thereby
JOURNAL Patent: WO 0190366-A 3483 29-NOV-2001;
Curagen Corporation (US)
FEATURES
source Location/Qualifiers
1..318
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 87.2 Length: 318
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x AX310498 (1-318)
Qy 1 GlyTyr***ValGluGlu 6
Db 165 GGATATTCTGTGAAGAA 182

RESULT 3
I08256 Sequence 1 from Patent EP 0369567. linear PAT 02-DEC-1994
LOCUS 330 bp DNA
DEFINITION I08256
ACCESSION I08256
VERSION I08256.1 GI:589029
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 330)
AUTHORS Johnson, M.J
TITLE Chimeric antibodies directed against metal chelates
JOURNAL Patent: EP 0369567-A2 1 23-MAY-1990;
Location/Qualifiers
1..330
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:

Pred. No.: 90.9 Length: 330
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x I08256 (1-330)

Qy 1 GlyTyr***ValGluGlu 6

Db 293 GGGTATTCGGTGGAGGAA 310

RESULT 4

I08257 I08257 387 bp DNA linear PAT 02-DEC-1994

LOCUS Sequence 3 from Patent EP 0369567.

DEFINITION I08257

ACCESSION I08257

VERSION I08257.1 GI:589030

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 387)

AUTHORS Johnson,M.J.

TITLE Chimeric antibodies directed against metal chelates

JOURNAL Patent: EP 0369567-A2 3 23-MAY-1990;

FEATURES Location/Qualifiers

source 1..387

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 109 Length: 387
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x I08257 (1-387)

Qy 1 GlyTyr***ValGluGlu 6

Db 350 GGGTATTCGGTGGAGGAA 367

RESULT 5

G79421 G79421 388 bp DNA linear STS 06-SEP-2002

LOCUS S210P6193RA3.T0 BALB/cByJ Mus musculus STS genomic, sequence tagged

DEFINITION site.

ACCESSION G79421

VERSION G79421.1 GI:22730177

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE 1 (bases 1 to 388)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

JOURNAL Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

COMMENT Unpublished (2002)

Contact: Kerstin Lindblad-Toh

Whitehead Institute for Biomedical Research, Center for Genome

Research

320 Charles Street, Cambridge, MA 02141, USA

Tel: 6172521477

Fax: 6172580903

Email: kersli@genome.wi.mit.edu

Primer A: None

Primer B: None

STS size: 388

Protocol:

WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 129S1/SvimJ, C3H/HeJ, and BALB/cByJ. The WGS reads were placed uniquely on the MGSCv3 C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated

as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

FEATURES

source

1..388

Location/Qualifiers

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="BALB/cByJ"

/db_xref="taxon:10090"

/map="12 22-343 63871472-63871794"

/clone_lib="BALB/cByJ"

<1..>388

STS

ORIGIN

Alignment Scores:
Pred. No.: 109 Length: 388
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 11 Gaps: 0

US-10-030-194A-5 (1-6) x G79421 (1-388)

Qy 1 GlyTyr***ValGluGlu 6

Db 238 GGATATAGTGTGGAGGAG 255

RESULT 6

CQ475389 CQ475389 400 bp DNA linear PAT 30-JAN-2004

LOCUS Sequence 7256 from Patent WO0160860.

DEFINITION CQ475389

ACCESSION CQ475389

VERSION CQ475389.1 GI:41441008

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE 1

AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.

TITLE Genes differentially expressed in human prostate cancer and their

JOURNAL use

Patent: WO 0160860-A 7256 23-AUG-2001;

Millennium Predictive Medicine, Inc. (US)

FEATURES Location/Qualifiers

source 1..400

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 113 Length: 400

Score: 27.00 Matches: 5

Percent Similarity: 83.33% Conservative: 0

Best Local Similarity: 83.33% Mismatches: 1

Query Match: 96.43% Indels: 0

DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x CQ475389 (1-400)

Qy 1 GlyTyr***ValGluGlu 6

Db 25 GGATATGCAGTAGAGGAA 42

RESULT 7
CQ505357
LOCUS CQ505357 431 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 37224 from Patent WO0160860.
ACCESSION CQ505357
VERSION CQ505357.1 GI:41471445
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 37224 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source
1..431
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 123 Length: 431
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0
US-10-030-194A-5 (1-6) x CQ505357 (1-431)
Qy 1 GlyTyr***ValGluGlu 6
Db 55 GGATATGCAGTAGAGGAA 72
RESULT 8
AX435885
LOCUS AX435885 479 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 4300 from Patent WO0229113.
ACCESSION AX435885
VERSION AX435885.1 GI:21660693
KEYWORDS Bacillus licheniformis
SOURCE Bacillus licheniformis
ORGANISM Bacillus licheniformis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Berka,R. and Clausen,I.G.
TITLE Methods for monitoring multiple gene expression
JOURNAL Patent: WO 0229113-A 4300 11-APR-2002;
Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)
FEATURES
source
1..479
/organism="Bacillus licheniformis"
/mol_type="unassigned DNA"
/db_xref="taxon:1402"
ORIGIN
Alignment Scores:
Pred. No.: 139 Length: 479
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0
US-10-030-194A-5 (1-6) x AX435885 (1-479)
Qy 1 GlyTyr***ValGluGlu 6
Db 334 GGATACAGTGTGGAAGAG 351

RESULT 9
AX306056
LOCUS AX306056 490 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 807 from Patent WO0188188.
ACCESSION AX306056
VERSION AX306056.1 GI:17645383
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 807 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
source
1..490
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.: 142 Length: 490
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0
US-10-030-194A-5 (1-6) x AX306056 (1-490)
Qy 1 GlyTyr***ValGluGlu 6
Db 344 GGCTATGCTGTGGAAGAG 361
RESULT 10
AX400752/c
LOCUS AX400752 498 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 428 from Patent WO0210453.
ACCESSION AX400752
VERSION AX400752.1 GI:21336932
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Mendrick,D., Porter,M.W., Johnson,K.R., Castle,A.L. and Elashoff,M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 428 07-FEB-2002;
Gene Logic, Inc. (US)
FEATURES
source
1..498
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="EMBL/GenBank Accession No. AA964455"
ORIGIN
Alignment Scores:
Pred. No.: 145 Length: 498
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0
US-10-030-194A-5 (1-6) x AX400752 (1-498)

```

QY      1 GlyTyr***ValGluGlu 6
Db      245 GGCTACTCTGTGTAAGAG 228

RESULT 11
LOCUS   AY584590          503 bp      RNA      linear      VRL 13-JUL-2004
DEFINITION   Nootka lupine vein-clearing virus RNA-dependent RNA polymerase
              gene, partial cds.
ACCESSION   AY584590
VERSION     AY584590.1  GI:50059158
KEYWORDS    Nootka lupine vein-clearing virus
SOURCE      Nootka lupine vein-clearing virus
ORGANISM    Viruses; ssRNA positive-strand viruses, no DNA stage;
              Tombusviridae.
REFERENCE   1 (bases 1 to 503)
AUTHORS    Robertson,N.L.
TITLE      The biology of a new virus isolated from Lupinus nootkatensis
           plants in Alaska
JOURNAL    Plant Pathol. (2004) In press
REFERENCE   2 (bases 1 to 503)
AUTHORS    Robertson,N.L.
TITLE      Direct Submission
JOURNAL    Submitted (26-MAR-2004) Arctic Plant Germplasm Introduction and
           Research Project, USDA, ARS, 533 E. Fireweed Ave., Palmer, AK
           99645, USA

FEATURES             Location/Qualifiers
     source           1..503
                     /organism="Nootka lupine vein-clearing virus"
                     /virus
                     /mol_type="genomic RNA"
                     /specific_host="Lupinus nootkatensis"
                     /db_xref="taxon:283876"
     CDS              <1..>503
                     /codon_start=3
                     /product="RNA-dependent RNA polymerase"
                     /protein_id="AA169242.1"
                     /db_xref="GI:50059159"
                     /translation="PLOSFFNRGTPDITLKLGLVKKPEKHACRALDRIMGPTVMKG
                     YTVREVGHIGRAGWCQSRPVAIGFDMRFPQHSVPALQFEHSLYMGCFNDFELRE
                     LLDWLQHSFGTGLSSNGYCRYKTEGGRKSDMTGLNCALACILTRHVLRGVPCRLI
                     NNGDDCV"

ORIGIN
Alignment Scores:      147          Length:      503
Pred. No.:            27.00        Matches:      5
Score:                83.33%       Conservative: 0
Percent Similarity:   83.33%       Mismatches:  1
Best Local Similarity: 83.33%       Indels:      0
Query Match:         96.43%        Gaps:        0
DB:                  14

US-10-030-194A-5 (1-6) x AY584590 (1-503)

QY      1 GlyTyr***ValGluGlu 6
Db      132 GGATATACCGTGAAGAA 149

RESULT 12
LOCUS   AX432800          535 bp      DNA      linear      PAT 28-JUN-2002
DEFINITION   Sequence 1215 from Patent WO0229113.
ACCESSION   AX432800
VERSION     AX432800.1  GI:21657604
KEYWORDS    Bacillus licheniformis
SOURCE      Bacillus licheniformis
ORGANISM    Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1
AUTHORS    Berka,R. and Clausen,I.G.
TITLE      Methods for monitoring multiple gene expression

```

```

JOURNAL    Patent: WO 0229113-A 1215 11-APR-2002;
FEATURES    Novozymes Biotech, Inc.: (US); Novozymes A/S (DK)
            Location/Qualifiers
            source
            1..535
            /organism="Bacillus licheniformis"
            /mol_type="unassigned DNA"
            /db_xref="taxon:1402"

ORIGIN
Alignment Scores:      157          Length:      535
Pred. No.:            27.00        Matches:      5
Score:                83.33%       Conservative: 0
Percent Similarity:   83.33%       Mismatches:  1
Best Local Similarity: 83.33%       Indels:      0
Query Match:         96.43%        Gaps:        0
DB:                  6

US-10-030-194A-5 (1-6) x AX432800 (1-535)

QY      1 GlyTyr***ValGluGlu 6
Db      373 GGATATTCAGTGAAGAA 390

RESULT 13
LOCUS   CQ527229          565 bp      DNA      linear      PAT 30-JAN-2004
DEFINITION   Sequence 59096 from Patent WO0160860.
ACCESSION   CQ527229
VERSION     CQ527229.1  GI:41493493
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS    Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE      Genes differentially expressed in human prostate cancer and their
           use
JOURNAL    Patent: WO 0160860-A 59096 23-AUG-2001;
           Millennium Predictive Medicine, Inc. (US)
FEATURES    Location/Qualifiers
            source
            1..565
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:      167          Length:      565
Pred. No.:            27.00        Matches:      5
Score:                83.33%       Conservative: 0
Percent Similarity:   83.33%       Mismatches:  1
Best Local Similarity: 83.33%       Indels:      0
Query Match:         96.43%        Gaps:        0
DB:                  6

US-10-030-194A-5 (1-6) x CQ527229 (1-565)

QY      1 GlyTyr***ValGluGlu 6
Db      209 GGATATGCAGTAGAGAA 226

RESULT 14
LOCUS   BV074683/c        582 bp      DNA      linear      STS 31-MAY-2003
DEFINITION   S209P6108RC7.T0 C3H/HeJ Mus musculus STS genomic, sequence tagged
           site.
ACCESSION   BV074683
VERSION     BV074683.1  GI:31190478
KEYWORDS    STS.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   1
AUTHORS    Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
           Eukaryota; Metazoa; Chordata; Sciurognathi; Muridae; Murinae; Mus.

```

REFERENCE 1 (bases 1 to 582)
AUTHORS Wade, C.M., Kulbokas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C., Lander, E.S., Lindblad-Toh, K. and Daly, M.J.
TITLE The mosaic structure of variation in the laboratory mouse genome
JOURNAL Nature 420 (6915), 574-578 (2002)
MEDLINE 22354684
COMMENT 12466852

Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 582

Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS reads were placed uniquely on the MGSVC3, C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated

as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

FEATURES
source
Location/Qualifiers
1..582
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C3H/HeJ"
/db_xref="taxon:10090"
/map="13 22-529 34703012-34703519"
/clone_lib="C3H/HeJ"
<1..>582

STS
ORIGIN
Alignment Scores:
Pred. No.: 173 Length: 582
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 11 Gaps: 0

US-10-030-194A-5 (1-6) x BV074683 (1-582)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 252 GGCTATGCTGTGGAGGAG 235

RESULT 15
G93830/c
LOCUS G93830 622 bp DNA linear STS 06-SEP-2002
DEFINITION S209P6221PH5.T0 C3H/HeJ Mus musculus STS genomic, sequence tagged site.
ACCESSION G93830
VERSION G93830.1 GI:22744586
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 622)
Wade, C.
Polymorphism Structure in the Mouse
Unpublished (2002)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome Research
320 Charles Street, Cambridge, MA 02141, USA

Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 622

Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS reads were placed uniquely on the MGSVC3, C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated

as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

FEATURES
source
Location/Qualifiers
1..622
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C3H/HeJ"
/db_xref="taxon:10090"
/map="13 22-575 30706949-30707504"
/clone_lib="C3H/HeJ"
<1..>622

STS
ORIGIN
Alignment Scores:
Pred. No.: 187 Length: 622
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 11 Gaps: 0

US-10-030-194A-5 (1-6) x G93830 (1-622)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 504 GGCTATGCTGTGGAGAA 487

RESULT 16
G86206/c
LOCUS G86206 623 bp DNA linear STS 06-SEP-2002
DEFINITION S208P6399RE4.T0 129S1/SvImJ Mus musculus STS genomic, sequence tagged site.
ACCESSION G86206
VERSION G86206.1 GI:22736962
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 623)
Wade, C.
Polymorphism Structure in the Mouse
Unpublished (2002)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 623

Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS reads were placed uniquely on the MGSVC3, C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated

as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

FEATURES

source
1. .623
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S1/SvimJ"
/db_xref="taxon:10090"
/map="9 22-573 74750233-74750779"
/clone.lib="129S1/SvimJ"
<1. .>623 .

STS
ORIGIN

Alignment Scores:
Pred. No.: 187 Length: 623
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 11 Gaps: 0

US-10-030-194A-5 (1-6) x G86206 (1-623)

Qy 1 GlyTyr***ValGluGlu 6
Db 58 GGATACTCTGTGAAGAA 41
|||||

RESULT 17
AF399524/c

LOCUS AF399524 650 bp DNA linear PRI 27-AUG-2001
DEFINITION Homo sapiens clone OR5D16 olfactory receptor gene, partial cds.

ACCESSION AF399524

VERSION AF399524.1 GI:15293632

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 650)

AUTHORS Fuchs,T., Malecova,B., Linhart,C., Sharan,R., Khen,M., Herwig,R., Shmulevich,D., Elkon,R., Steinfach,M., O'Brien,J.K., Radelof,U., Leirach,H., Olender,Z., Glusman,G., Lancet,D. and Shamir,R.

DEFOG: A Practical Scheme for Deciphering Families of Genes

Unpublished

2 (bases 1 to 650)

Fuchs,T., Malecova,B., Linhart,C., Sharan,R., Khen,M., Herwig,R., Shmulevich,D., Elkon,R., Steinfach,M., O'Brien,J.K., Radelof,U., Leirach,H., Olender,Z., Glusman,G., Lancet,D. and Shamir,R.

Direct Submission

TITLE Submitted (16-JUL-2001) Dept. of Molecular Genetics and the Crown Human Genome Center, Weizmann Institute of Science, Rehovot 76100, Israel

FEATURES
source

1. .650
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="51.267026"
/clone="OR5D16"
<1. .>650
/product="olfactory receptor"
<1. .>650
/note="contains transmembrane regions 2-7"
/codon_start=1
/product="Olfactory receptor"
/protein_id="AAK95009.1"
/db_xref="GI:15293633"
/translation="FVDFECYSIIAPMVLNVLVVEDRTISPGCLVFFPFECTVWTE
LILFAVWAYDFHVAICPLLYTVAISQKLCAMLVVLYVANGVACSLILACSLKLSFH
GFNTINHFCELSLISLSPDSYLSLFTVATFNEISTLLILTSYAFIIVTLK
MPSASGHRKVFSTCASHLTATIFHGTILFLYCVFNKSNRHTVKVASVFYTVIP"

mRNA

CDS

ORIGIN

Alignment Scores:
Pred. No.: 196 Length: 650
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-5 (1-6) x AF399524 (1-650)

Qy 1 GlyTyr***ValGluGlu 6
Db 585 GGGTACACAGTAGAGAA 568
|||||

RESULT 18
AV170535

LOCUS AV170535 662 bp mRNA linear ROD 01-OCT-2003
DEFINITION Mus musculus clone ll4s62 immunoglobulin lambda chain mRNA, partial cds.

ACCESSION AV170535

VERSION AV170535.1 GI:37221132

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 662)

AUTHORS Mao,C., Zhang,X., Puthenveetil,M., Carroll,M.C. and

Imanishi-Kari,I.
T Cell-Independent Somatic Hypermutation in Murine B Cells with an Immature Phenotype

Unpublished

2 (bases 1 to 662)

Mao,C., Zhang,X., Puthenveetil,M., Carroll,M.C. and

Imanishi-Kari,I.

Direct Submission

TITLE Submitted (30-OCT-2002) Department of Pathology and Program in Immunology, Tufts University School of Medicine and Sackler School of Graduate Biomedical Sciences, 136 Harrison Ave, Boston, MA 02111, USA

FEATURES
source

1. .662
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="114862"
/cell_type="493+lambda+"
/tissue_type="bone marrow"
/note="quasi-monoclonal (QM); C4-/-; CD3epsilon-/-"
1. .>662
/codon_start=1
/product="immunoglobulin lambda chain"
/protein_id="AAO53378.1"
/db_xref="GI:37221133"
/translation="MAWISLILSLALSSGAISQAVVTQESALTSPGETVLTCSRSS
TGAVTASNANWVQEKPDHLFTLIGTNNRAPGVAPRFGSLIGDKAALTITGAOTE
DEALYFCALVYNHWPVGGTKLTVLGPKSPSVTLFPSPSSELETNKATLVCTITD
FYGVVTVDWKVDGTPVTQGMETTPQSKSNKNKYNASSYLTLTARAWERHSSYSCQVT
HEG"
1. .57
sig_peptide
V_region 58. .348
/note="VLI"
J_segment 349. .385
/note="J1"

ORIGIN

Alignment Scores:
Pred. No.: 200 Length: 662
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1

```
Query Match: 96.43% Indels: 0
DB: 10 Gaps: 0

US-10-030-194A-5 (1-6) x AY170535 (1-662)

Qy 1 GlyTyr***ValGluGlu 6
Db 350 GGGTATTTCGGTGGAGAA 367

RESULT 19
BV042227/c
LOCUS BV042227 681 bp DNA linear STS 31-MAY-2003
DEFINITION S212P6005FH11.T0 CZECHII/Ei Mus musculus STS genomic, sequence
tagged site.
ACCESSION BV042227
VERSION BV042227.1 GI:31126122
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 681)
AUTHORS Wade, C.M., Kulbokas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C.,
Lander, E.S., Lindblad-Toh, K. and Daly, M.J.
TITLE The mosaic structure of variation in the laboratory mouse genome
JOURNAL Nature 420 (6915), 574-578 (2002)
MEDLINE 22354684
PubMed 12466852
COMMENT
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 681
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvimJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGSVC3 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.

FEATURES
source
1..681
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CZECHII/Ei"
/db_xref="taxon:10090"
/map="- 1 22-610 25146058-25145452"
/clone_lib="CZECHII/Ei"
<1..>681

STS
ORIGIN
Alignment Scores:
Pred. No.: 207 Length: 681
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 11 Gaps: 0

US-10-030-194A-5 (1-6) x BV042227 (1-681)

Qy 1 GlyTyr***ValGluGlu 6
Db 304 GGGTACACAGTTGAGAA 287
```

```
RESULT 20
BD214758/c
LOCUS BD214758 708 bp DNA linear PAT 17-JUL-2003
DEFINITION Novel human genes and gene expression products ii.
ACCESSION BD214758
VERSION BD214758.1 GI:33024528
KEYWORDS JP 2002519000-A/2900.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 708)
AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.S.,
Reinhard, C., Giese, K., Randazzo, F., Kennedy, G.C., Pot, D.,
Kassam, A., Lamson, G., Drmanac, R., Crkvenjakov, R., Dickson, M.,
Drmanac, S., Labat, I., Leshkowitz, D., Kita, D., Garcia, V., Jones, L.W.
and Crain, B.S.
TITLE Novel human genes and gene expression products ii
JOURNAL Patent: JP 2002519000-A 2900 02-JUL-2002;
CHIRON CORP. HYSQ INC
COMMENT OS Homo sapiens (human)
PN JP 2002519000-A/2900
PD 02-JUL-2002
PF 28-JAN-1999 JP 2000556580
PR 28-JUN-1998 US 60/072910, 24-FEB-1998 US 60/075954 PR
31-MAR-1998 US 60/080114, 03-APR-1998 US 60/080515 PR
03-APR-1998 US 60/080666, 21-OCT-1998 US 60/105234 PR
28-OCT-1998 US 60/105877
PI LOUIS T WILLIAMS, JAIME ESCOBEDO, MICHAEL A INNIS, PABLO PI
DOMINGUEZ GARCIA,
PI JULIE SUDDUTH KLINGER, CHRISTOPH REINHARD, KLAUSE GIESE, FILIPPO
PI RANDAZZO,
PI GIULIA C KENNEDY, DAVID POT, ALTAF KASSAM, GEORGE LAMSON, RADOJE
PI DRMANAC,
PI RADOMIR CRKVENJAKOV, MARK DICKSON, SNEZANA DRMANAC, IVAN LABAT,
PI DENA LESHKOWITZ, DAVID KITA, VERONICA GARCIA, LEE WILLIAM JONES,
PI BIRJIT STACHE CRAIN
PC C12N15/09, C12N15/09, C07K14/47, C07K14/82, C07K16/18, C12N1/15, PC
C12N1/19,
PC C12N1/21, C12N5/10, C12Q1/68, C12N15/00, C12N5/00, C12N15/00 CC n
= A, T, C or G
FH Key Location/Qualifiers
FT misc feature (1)..(708).
FEATURES
source
1..708
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 216 Length: 708
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x BD214758 (1-708)

Qy 1 GlyTyr***ValGluGlu 6
Db 413 GGTATAGTGTGGAAGAG 396

RESULT 21
AY102903/c
LOCUS AY102903 732 bp DNA linear BCT 24-JUN-2002
DEFINITION Uncultured bacterium clone BCM3S-13B 16S ribosomal RNA gene,
partial sequence.
ACCESSION AY102903
VERSION AY102903.1 GI:21552648
KEYWORDS uncultured bacterium
SOURCE uncultured bacterium
```

```

ORGANISM      uncultured bacterium
REFERENCE      Bacteria; environmental samples.
AUTHORS        1 (bases 1 to 732)
TITLE          Prenger,J.P., Castro,H. and Ogram,A.
               Diversity of Bacterial and Archaeal Communities Associated with
               Periphyton in a Subtropical Freshwater Marsh
JOURNAL        Unpublished in a Subtropical Freshwater Marsh
AUTHORS        2 (bases 1 to 732)
TITLE          Prenger,J.P., Castro,H. and Ogram,A.
               Direct Submission
JOURNAL        Submitted (11-MAY-2002) Soil and Water Science Department,
               University of Florida, P.O. Box 110510, 106 Newell Hall,
               Gainesville, FL 32611, USA
FEATURES       Location/Qualifiers
source         1..732
               /organism="uncultured bacterium"
               /mol_type="genomic DNA"
               /isolation_source="subtropical freshwater marsh"
               /db_xref="taxon:77133"
               /clone="BCW3S-13B"
               /environmental_sample
               <1..>732
               /product="16S ribosomal RNA"

rRNA

ORIGIN
Alignment Scores:
Pred. No.:      225      Length:      732
Score:          27.00    Matches:      5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match:    96.43% Indels:      0
DB:             1       Gaps:        0

US-10-030-194A-5 (1-6) x AY102903 (1-732)

Qy      1 GlyTyr***ValGluGlu 6
Db      127 GGCTATTCCTGGTAGAG 110

RESULT 22
LOCUS      AY662689      772 bp      mRNA      linear      VRT 24-JUL-2004
DEFINITION Gallus gallus curly protein (CURP) mRNA, partial cds.
ACCESSION  AY662689
VERSION     AY662689.1 GI:50380134
KEYWORDS   Gallus gallus (chicken)
SOURCE     Gallus gallus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
               Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 772)
AUTHORS    Pagano,A., Tonachini,L., Monticone,M., Tortelli,F., Randazzo,N.,
               Tavella,S., Di Marco,E., Cancedda,R. and Castagnola,P.
TITLE      CURP (curly protein) is a novel developmental regulated
               leucine-zipper protein containing a PDZ-ligand and a Y-based
               sorting signal
JOURNAL    Unpublished
AUTHORS    2 (bases 1 to 772)
TITLE      Pagano,A., Tonachini,L., Monticone,M., Tortelli,F., Randazzo,N.,
               Tavella,S., Di Marco,E., Cancedda,R. and Castagnola,P.
JOURNAL    Direct Submission
GENETICA, Universita' di Genova, Largo R.Benzi, 10, Genova, GE
16132, Italy
FEATURES       Location/Qualifiers
source         1..772
               /organism="Gallus gallus"
               /mol_type="mRNA"
               /db_xref="taxon:9031"
               <1..772
               /gene="CURP"
               <1..213
gene
CDS

```

```

/ gene="CURP"
/ note="leucin-zipper protein"
/ codon_start=1
/ product="curly protein"
/ protein_id="AAT76305.1"
/ db_xref="GI:50380135"
/ translation="LKSKDRALDREIAQLLSEGYSVELEKHSILLHEYNEIKDAGQM
L L G K L A V I R G V T T K Q L I P E Y D L E L N D"

ORIGIN
Alignment Scores:
Pred. No.:      239      Length:      772
Score:          27.00    Matches:      5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match:    96.43% Indels:      0
DB:             5       Gaps:        0

US-10-030-194A-5 (1-6) x AY662689 (1-772)

Qy      1 GlyTyr***ValGluGlu 6
Db      55 GGGTACAGCGTGGAGGAG 72

RESULT 23
LOCUS      AR346929      849 bp      DNA      linear      PAT 17-AUG-2003
DEFINITION Sequence 1540 from patent US 6583275.
ACCESSION  AR346929
VERSION     AR346929.1 GI:33743974
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 849)
AUTHORS    Doucette-Stamm,L.A. and Bush,D.
TITLE      Nucleic acid sequences and expression system relating to
               Enterococcus faecium for diagnostics and therapeutics
JOURNAL    Patent: US 6583275-A 1540 24-JUN-2003;
FEATURES    Location/Qualifiers
source      1..849
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      266      Length:      849
Score:          27.00    Matches:      5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match:    96.43% Indels:      0
DB:             6       Gaps:        0

US-10-030-194A-5 (1-6) x AR346929 (1-849)

Qy      1 GlyTyr***ValGluGlu 6
Db      694 GGATATAGTGTGGAAGAG 677

RESULT 24
LOCUS      BT012980      873 bp      mRNA      linear      PLN 11-MAY-2004
DEFINITION Lycopersicon esculentum clone 114192F, mRNA sequence.
ACCESSION  BT012980
VERSION     BT012980.1 GI:47104395
KEYWORDS   FLI_CDNA.
SOURCE     Lycopersicon esculentum (tomato)
ORGANISM   Lycopersicon esculentum
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE  1 (bases 1 to 873)
AUTHORS    Kirkness,E.F., Wang,W. and Vazeille,A.

```

TITLE
JOURNAL
FEATURES
source

Direct Submission
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
Location/Qualifiers
1. .873
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="114192P"
/tissue_type="root, plant at fruit set"
/note="TRXAH19"

ORIGIN

Alignment Scores: 274 Length: 873
Pred. No.: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x BT012980 (1-873)

Qy 1 GlyTyr***ValGluGlu 6
|||||
256 GGATATACAGTTGAGGAA 239

Db 256 GGATATACAGTTGAGGAA 239

RESULT 25
BC071487 926 bp mRNA linear VRT 02-JUN-2004
LOCUS
DEFINITION
Danio rerio cDNA clone MGC:86840 IMAGE:6902526, complete cds.
ACCESSION
BC071487
VERSION
BC071487.1 GI:47938022
KEYWORDS
MGC.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
AUTHORS
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F.,
Diatchenko,L., Narusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,J.L.,
Schaeetz,T.E., Brownstein,M.J., Udén,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalau,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

TITLE
JOURNAL
PUBMED
AUTHORS
Strausberg,R.
TITLE
JOURNAL
Submitted (01-JUN-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Genome Institute of Singapore

FEATURES
source

1. .926
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:86840 IMAGE:6902526"
/tissue_type="Embryo, 7 different stages (from just
fertilized embryos to 72 hours just hatched baby fish)"
/clone_lib="GISZF001"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
53..628
/codon_start=1
/product="Unknown (protein for MGC:86840)"
/protein_id="AAH71487.1"
/db_xref="GI:47938023"
/translation="MPVCVMAELDEKLLQFQTELEAPGGIATPOVYSOLLVYL
HNEMNARYLWKRIPOAIKANTPEMAAIWAGQRIWQDPGGIYSAIAVQWSSILP
VMEALRESTRRRYGLVAQYATISASDFAFVGYSVEEAVKGVVSHGWQADPNTRMI
MPQKDPDPVPSLVPEEQQLRLTQDYVAFLEN"

ORIGIN

Alignment Scores: 293 Length: 926
Pred. No.: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BC071487 (1-926)

Qy 1 GlyTyr***ValGluGlu 6
|||||
458 GGATACTCGGTAGAAGAG 475

Db 458 GGATACTCGGTAGAAGAG 475

RESULT 26
AX241744/c 963 bp DNA linear PAT 26-SEP-2001
LOCUS
DEFINITION
Sequence 492 from Patent WO0127158.
ACCESSION
AX241744
VERSION
AX241744.1 GI:15798619
KEYWORDS
synthetic construct
SOURCE
synthetic construct
artificial sequences.
ORGANISM
1
REFERENCE
AUTHORS
Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
Yanai,I.
TITLE
JOURNAL
Patent: WO 0127158-A 492 19-APR-2001;
Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES
source
1. .963
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

REMARK
COMMENT

cdNA Library Preparation: S. Mathavan, Chia-Lin Wei, and Yijun
Ruan, Genome Institute of Singapore
cdNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAL Plate: 56 Row: k Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

```

/notice="(H38g341 nucleotide)"

ORIGIN
Alignment Scores:
Pred. No.: 307 Length: 963
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x AX241744 (1-963)
Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 792 GGGTACACAGTAGAGGAA 775

RESULT 27
AK062617 964 bp mRNA linear PLN 24-JUL-2003
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:001-105-B04, full
DEFINITION insert sequence.
ACCESSION AK062617.1 GI:32972635
VERSION Full_CDNA; oligo-capping.
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
1 The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team:
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuka,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group, Ootomo,Y., Murakami,K.,
Ida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Ootato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M., and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
2 (bases 1 to 964)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Inamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuka,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Ootomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,

ORIGIN
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:shikuchien@affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cdna/
NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
Negata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuka,K., Shishiki,T., and
Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K., and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
Oka,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
Yasunishi,A., and Hayashizaki,Y.
Location/Qualifiers
1. 964
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clones="001-105-B04"

ORIGIN
Alignment Scores:
Pred. No.: 307 Length: 964
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x AK062617 (1-964)
Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 434 GGATACCGGTGGAGGAG 451

RESULT 28
AY267671 986 bp mRNA linear INV 02-JUL-2003
LOCUS Bigelowiella natans plastid protein SufE mRNA, partial cds; nuclear
DEFINITION Bigelowiella natans plastid product.
ACCESSION AY267671.1 GI:32307559
VERSION AY267671
KEYWORDS SOURCE
ORGANISM Bigelowiella natans
Eukaryota; Cercozoa; Chlorarachniophyceae; Bigelowiella.
REFERENCE
1 (bases 1 to 986)
AUTHORS Archibald,J.M., Rogers,M.B., Toop,M., Ishida,K. and Keeling,P.J.
Lateral gene transfer and the evolution of plastid-targeted
TITLE

```

proteins in the secondary plastid-containing alga *Bigelowiella natans*
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7678-7683 (2003)

JOURNAL MEDLINE PUBMED
22709102 12777624
REFERENCE 2 (bases 1 to 986)
AUTHORS Archibald, J.M., Rogers, M.B., Toop, M., Ishida, K. and Keeling, P.J.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2003) Department of Botany, University of British Columbia, 3529-6270 University Blvd., Vancouver, BC V6T 1Z4, Canada

FEATURES
source Location/Qualifiers
1..986 /organism="Bigelowiella natans"
/mol_type="mRNA"
/strain="CCMP 621"
/db_xref="taxon:227086"
<1..700 /codon_start=2
CDS /product="plastid protein Sufe"
/protein_id="AAP79185.1"
/db_xref="GI:32307560"
/translations="TPYHFGASLLLVVAVASSIALSPTRHLSTGLKPSMVLQRPOPTGVLCQVTTRKLEVSRLNAGRSFRAMEQLSRSSRQREMHAFATTTSSSLNLTPNLEKEVRLKSADPNTLRHQOLLYAKEAQLPSQKAENKVPGLSTVVVAVAKDDGKVFFRGSDALITKLGNLLIRLGLSGYSVEEIVAVKPFIQEAGITQSITPCRNNGFLNM LNTMKCKAQVQTR"

ORIGIN
Alignment Scores:
Pred. No.: 315 Length: 986
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x AY267671 (1-986)

Qy 1 GLYTYR***ValGluGlu 6
||| ||| ||| ||| ||| |||
Db 551 GGATATTTCAGTGAAGAG 568

RESULT 29
AX318263/c AX318263 987 bp DNA linear PAT 14-DEC-2001
LOCUS Sequence 44 from Patent WO0190359.
DEFINITION AX318263
ACCESSION AX318263
VERSION AX318263.1 GI:17900923
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Patterson, C., Tribouley, C.M., Yao, M.G., Griffin, J.A., Thornton, M., Lu, Y., Kallick, D.A., Gandhi, A.R. and Au-Young, J.
TITLE G-protein coupled receptors
JOURNAL Patent: WO 0190359-A 44 29-NOV-2001;
Incyte Genomics, Inc. (US)

FEATURES
source Location/Qualifiers
1..987 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7476117CB1"

ORIGIN
Alignment Scores:
Pred. No.: 315 Length: 987
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0

```

PN JP 2002112793-A/181
PD 16-APR-2002
PF 09-FEB-2001 JP 2001034434
PI TATSUYA HAGA, SHIGEKI TAKEDA, NARIKI MIYAKE
PC
C12N15/09, A01K67/027, A61K38/00, A61K39/395, A61K39/395, A61K45/00, PC
A61K48/00,
C12N1/21,
PC A61P43/00, C07K14/705, C07K16/28, C07K19/00, C12N1/15, C12N1/19, PC
C12N5/10, C12Q1/02, C12Q1/68, G01N33/15, G01N33/50, G01N33/53, PC
G01N33/566//
PC C12N5/10, C12N15/00, A61K37/02, C12N5/00
CC Novel G-protein coupled receptors
FH Key Location/Qualifiers
FT CDS (1)..(987)

FEATURES
    source
        1..987
            Location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 315 Length: 987
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x BD144456 (1-987)

QY 1 GlyTyr***ValGluGlu 6
|||||
DB 792 GGGTACACAGTAGAGGAA 775

RESULT 32
AX556236/c
LOCUS AX556236 1002 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 31 from Patent WO0250117.
ACCESSION AX556236
VERSION AX556236.1 GI:25899588
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Padigaru, M., Kekuda, R., Colman, S.D., Spytek, K.A., Ballinger, R.A.,
Vernet, C.A., Li, L., Shenoy, S. and Casman, S.J.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0250117-A 31 27-JUN-2002;
Curagen Corporation (US)
FEATURES
    source
        1..1002
            Location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 321 Length: 1002
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x AX556236 (1-1002)

QY 1 GlyTyr***ValGluGlu 6
|||||
DB 798 GGGTACACAGTAGAGGAA 781

RESULT 33
AY659169
LOCUS AY659169 1035 bp DNA linear SYN 07-JUL-2004
DEFINITION Synthetic construct Pseudomonas aeruginosa clone FLH036922.01F
ACCESSION AY659169
VERSION AY659169.1 GI:49087329
KEYWORDS Pseudomonas aeruginosa ORF Project.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 1035)
AUTHORS Qiu, Q., Anumanthan, A., Mar, W., Zuo, D., Murthy, T.V.S., Hu, Y.,
Taycher, E., Halleck, A., Hainsworth, E., Lory, S., LaBaer, J. and
Brizuela, L.
TITLE The Pseudomonas aeruginosa gene repository
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1035)
AUTHORS Qiu, Q., Anumanthan, A., Mar, W., Zuo, D., Murthy, T.V.S., Hu, Y.,
Taycher, E., Halleck, A., Hainsworth, E., Lory, S., LaBaer, J. and
Brizuela, L.
DIRECT SUBMISSION
SUBMITTED (15-JUN-2004) Biological Chemistry and Molecular
Pharmacology, Harvard Institute of Proteomics, Harvard Institute of
Proteomics, 320 Charles St., Cambridge, MA 02141, USA
This clone is part of the Pseudomonas aeruginosa PA01 ORF clone
collection generated by the Harvard Institute of Proteomics. Each
CDS has been PCR-amplified from genomic DNA. Initiation codon has
been normalized to ATG and stop codon has been substituted for GGA
to allow generation of C-terminal tagged proteins. AttB
recombination sites have been added to either end of the orf and
directionally cloned using the Gateway cloning system into pDONR
201. Additional sequences in the clone: 'GAAGGAGATACC' before the
'ATG' (corresponding to ribosomal binding site and Kozak consensus
sequences).
FEATURES
    Location/Qualifiers
        1..1035
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
            /clones="FLH036922.01F"
            /lab_host="Escherichia coli DH5alpha T1 resistant"
            /note="derived from Pseudomonas aeruginosa"
        1..>1035
            /codon_start=1
            /transl_table=11
            /product="PA0321"
            /protein_id="AAT51444.1"
            /db_xref="GI:49087330"
            /translation="MLTIYSDHRLHHRHLEIGGQFTPCFEKFSRADMVLDVRKAVG
LGEVRAPDFGLEPIRRVHSEGFVFLQAWQDWLATGRSHDMLPIAMPTRLRLOTEP
LDNIDRLGYVSFDAGAPITAGTWOAITSSANVALSGOSELANGARSVSLCRPGGHA
AADYNGGCGCFNNAIAAQAFLDGRAGRVALLDVDYHHGNGTQDIFDPRADVLFVSIH
GDPRFEYPFLGYADEKNGVGTYNPNPLAGSDWATWSQAQAALRQIQAYANDA
LIVSLGVDTPFKGDPISQFRLDSPDYLRMGEAIGKLGATLFWMEGGYAVEBIGINAVN
VLQGFEGVHRG"
        misc_difference 863
        misc_difference 1033..1034
            /replace="a"
            /replace="ta"

ORIGIN
Alignment Scores:
Pred. No.: 333 Length: 1035
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 12 Gaps: 0

US-10-030-194A-5 (1-6) x AY659169 (1-1035)

```

```

Oy 1 GlyTyr***ValGluGlu 6
    ||||| ||||| ||||| |||||
Db 964 GGCTACGGCGTGGAGGAA 981

RESULT 34
CR523161
LOCUS Cucurbita maxima Pufg mRNA linear VRT 15-JUN-2004
DEFINITION Gallus gallus finished cDNA, clone CHEST319124.
ACCESSION CR523161
VERSION CR523161.1 GI:48841728
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1061)
AUTHORS Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
Tickle,C. and Wilson,S.A.
TITLE Direct Submission
JOURNAL Submitted (15-JUN-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickens@hms.umist.ac.uk
COMMENT BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
sequencing project.
This sequence is from the
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from heart, normalised, and poly A-tailed.
EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
coli DH10B.

FEATURES
source 1..1061
    /organism="Gallus gallus"
    /mol_type="mRNA"
    /strain="White Leghorn, Hisex"
    /db_xref="taxon:9031"
    /clone="CHEST319124"
    /clone_lib="CSEQCHN71"
    /dev_stage="stage 36"

ORIGIN
Alignment Scores: 342 Length: 1061
Pred. No.: 27.00 Matches: 5
Score: 83.33% Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.43% Gaps: 0
DB: 5

US-10-030-194a-5 (1-6) x CR523161 (1-1061)

Oy 1 GlyTyr***ValGluGlu 6
    ||||| ||||| ||||| |||||
Db 261 GGGTACACGGTGGAGGAG 278

RESULT 35
AB059484
LOCUS Cucurbita maxima Pufg mRNA linear PLN 12-SEP-2002
DEFINITION Cucurbita maxima (winter squash)
ACCESSION AB059484
VERSION AB059484.1 GI:22830572
KEYWORDS
SOURCE Cucurbita maxima
ORGANISM Cucurbita maxima
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
REFERENCE 1

Fujita,M. and Hossain,M.Z.
Purification of a phi-type glutathione S-transferase from pumpkin
flowers, and molecular cloning of its cDNA
Biosci. Biotechnol. Biochem. 66, 2068-2075 (2002)
2 (bases 1 to 1064)
Direct Submission
Fujita,M. and Hossain,M.Z.
Submitted (07-APR-2001) Masayuki Fujita, Kagawa University, Faculty
of Agriculture; 2393 Ikenobe, Miki-cho, Kita-gun, Kagawa 761-0795,
Japan (E-mail:fujita@ag.kagawa-u.ac.jp, Tel:81-087-891-3133,
Fax:81-087-891-3021)
Japan (E-mail:fujita@ag.kagawa-u.ac.jp, Tel:81-087-891-3133,
Fax:81-087-891-3021)

FEATURES
source 1..1064
    /organism="Cucurbita maxima"
    /mol_type="mRNA"
    /db_xref="taxon:3661"
    /gene="pugf"
    /gene="pugf"
    /codon_start=1
    /product="glutathione S-transferase"
    /protein_id="BAC15623.1"
    /db_xref="GI:22830573"
    /translation="MAGIKVHGIPSTATARVLAALYKGLQFELVNVKLEHGHKKE
PFISINFGQVPGFNGDGLTLFESRAITQVISGNVSNSTELIPKDSKKGAVLTWIE
VESHRDPVAMKLVFEFLKPLGLGDADPAVVEKSEADLQGLDIIYEKRLSGSKYLG
GDSFSLADLHLPLVIGLYLLATPSKLFESRPHVNAVVGDIARPSWAKVLARK"

ORIGIN
Alignment Scores: 344 Length: 1064
Pred. No.: 27.00 Matches: 5
Score: 83.33% Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 96.43% Indels: 0
Query Match: 96.43% Gaps: 0
DB: 8

US-10-030-194a-5 (1-6) x AB059484 (1-1064)

Oy 1 GlyTyr***ValGluGlu 6
    ||||| ||||| ||||| |||||
Db 587 GGCTACACCGTGGAGGAA 604

RESULT 36
AR346914
LOCUS AR346914 1065 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 1525 from patent US 6583275.
ACCESSION AR346914
VERSION AR346914.1 GI:33743959
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1065)
AUTHORS Doucette-Stamm,L.A. and Bush,D.
TITLE Nucleic acid sequences and expression system relating to
Enterococcus faecium for diagnostics and therapeutics
JOURNAL Patent: US 6583275-A 1525 24-JUN-2003;
FEATURES
source 1..1065
    /organism="unknown"
    /mol_type="genomic DNA"

ORIGIN
Alignment Scores: 344 Length: 1065
Pred. No.: 27.00 Matches: 5
Score: 83.33% Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 96.43% Indels: 0
Query Match: 96.43% Gaps: 0
DB: 6

```

[illegible]

```
COMMENT
CB10 ISA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
BSBRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus cdna
sequencing project.
This sequence is from the
BSBRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from heads, normalised, and poly A-trimmed.
ECORI-NotI cut cDNA was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
coli DH10B.

FEATURES
    source
    1..1203
    /organism="Gallus gallus"
    /mol_type="mRNA"
    /strain="White Leghorn, Hisex"
    /db_xref="taxon:9031"
    /clone="CHEST766c12"
    /clone_lib="CSEQCHN23"
    /dev_stage="stage 22"

ORIGIN
Alignment Scores:
Pred. No.: 395 Length: 1203
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BX930000 (1-1203)
Qy 1 GlyTyr***ValGluGlu 6
||||| ||||| ||||| |||||
Db 22 GGGTACACGGTGGAGGAG 39

RESULT 40
LOCUS CR352483 1214 bp mRNA linear VRT 11-MAR-2004
DEFINITION Gallus gallus finished cDNA, clone CHEST17917.
ACCESSION CR352483
VERSION CR352483.1 GI:45422089
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
Tickle,C. and Wilson,S.A.
Direct Submission
Submitted (09-MAR-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
On Apr 1, 2004 this sequence version replaced gi:41634457.
BSBRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus cdna
sequencing project.
This sequence is from the
BSBRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from cerebrum, and poly A-trimmed. EORI-NotI
cut cDNA was then ligated into the vector. Vector: pBluescript II
KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia coli DH10B.

FEATURES
    source
    1..1214
    /organism="Gallus gallus"
    /mol_type="mRNA"
    /strain="Compton Line 151"
    /db_xref="taxon:9031"
    /clone="CHEST13017"
    /clone_lib="CSEQCHL15"
    /dev_stage="adult"

ORIGIN
Alignment Scores:
Pred. No.: 401 Length: 1220
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BX933929 (1-1220)
Qy 1 GlyTyr***ValGluGlu 6
||||| ||||| ||||| |||||
Db 39 GGGTACACGGTGGAGGAG 56

RESULT 42
LOCUS CR352483 1214 bp mRNA linear VRT 30-MAR-2004
DEFINITION Gallus gallus finished cDNA, clone CHEST13017.
ACCESSION CR352483
VERSION CR352483.2 GI:46018039
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
Tickle,C. and Wilson,S.A.
Direct Submission
Submitted (29-MAR-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
On Apr 1, 2004 this sequence version replaced gi:41634457.
BSBRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus cdna
sequencing project.
This sequence is from the
BSBRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from cerebrum, and poly A-trimmed. EORI-NotI
cut cDNA was then ligated into the vector. Vector: pBluescript II
KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia coli DH10B.

FEATURES
    source
    1..1220
    /organism="Gallus gallus"
    /mol_type="mRNA"
    /strain="Compton Line 151"
    /db_xref="taxon:9031"
    /clone="CHEST13017"
    /clone_lib="CSEQCHL15"
    /dev_stage="adult"

ORIGIN
Alignment Scores:
Pred. No.: 401 Length: 1220
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BX933929 (1-1220)
Qy 1 GlyTyr***ValGluGlu 6
||||| ||||| ||||| |||||
Db 39 GGGTACACGGTGGAGGAG 56

RESULT 42
```

```

S55498
LOCUS       S55498             1221 bp    DNA        linear    INV 08-NOV-1996
DEFINITION   G alpha 4-Guanine nucleotide-binding protein [Dictyostelium sp.,
ACCESSION    Genomic, 1221 nt].
VERSION      S55498
KEYWORDS     S55498.1 GI:235477
SOURCE       Dictyostelium sp.
ORGANISM     Dictyostelium sp.
REFERENCE    1 (bases 1 to 1221)
AUTHORS      Hadwiger, J.A., Wilkie, T.M., Strathmann, M. and Firtel, R.A.
TITLE        Identification of Dictyostelium G alpha genes expressed during
multicellular development
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 88 (18), 8213-8217 (1991)
MEDLINE      91376114
PUBMED       1910174
REMARK       GenBank staff at the National Library of Medicine created this
entry [NCBI gisbseq 5498] from the original journal article.
FEATURES             Location/Qualifiers
     source          1..1221
                     /organism="Dictyostelium sp."
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5784"
     gene            1..1221
                     /gene="G alpha 4"
     CDS             112..1149
                     /note="Guanine nucleotide-binding protein"
                     /gene="G alpha 4"
                     /note="Guanine nucleotide-binding protein"
                     /codon_start=1
                     /protein_id="AAB19793.1"
                     /db_xref="GI:235478"
                     /translation="MRPKFGSEETEQQSKIDKSIETDRRKLKDVKKLLILPGESGK
STIPKMKIIODGGSVLELLEAFVYVNCISOMKALLTASAKINTELEVENKORA
ANLVPRRTGNSEPMILLADIKHLMDKCIKETYAKDKHFNLDNSAAYFFENIDRYR
EDFVNEQULRCRVRTTIGIOSEFTFKIRLKYDVVGQSRQRKWIHCFDCVTAVI
FVAAMSDYDVLREDSVNRRESIALPFIKNCIDYFKETPIVLFLNKKDLFKKLLKR
FVLAQSCFSDYTGPNKYKDAAMFIQSQYLAQGPSRTIYTHATCAVDTENIKFVFAVR
QTILSQALEHF"
ORIGIN
Alignment Scores:
Pred. No.:      402          Length:      1221
Score:          27.00        Matches:      5
Percent Similarity: 83.33%   Conservative: 0
Best Local Similarity: 83.33% Mismatches:  1
Query Match:    96.43%      Indels:      0
DB:             3           Gaps:        0

US-10-030-194A-5 (1-6) x S55498 (1-1221)

Qy      1  GlyTyr***ValGluGlu 6
      ||||| ||||| ||||| ||||| |||||
Db      286 GGTATAGTGTGAAGAA 303

RESULT 43
LOCUS       CQ645844             1239 bp    DNA        linear    PAT 02-FEB-2004
DEFINITION   Sequence 2801 from Patent WO0234771.
ACCESSION    CQ645844
VERSION      CQ645844.1 GI:41682648
KEYWORDS     Streptococcus agalactiae
SOURCE       Streptococcus agalactiae
ORGANISM     Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE    1
AUTHORS      Telford, J., Massignani, V., Margarit y Ros, I., Grandi, G., Fraser, C.
and Tettelin, H.
TITLE        Nucleic acids and proteins from streptococcus groups a & b
JOURNAL      Patent: WO 0234771-A 2801 02-MAY-2002;
Chiron s.p.a. (IT) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)

```

```

FEATURES             Location/Qualifiers
     source          1..1239
                     /organism="Streptococcus agalactiae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1311"
ORIGIN
Alignment Scores:
Pred. No.:      408          Length:      1239
Score:          27.00        Matches:      5
Percent Similarity: 83.33%   Conservative: 0
Best Local Similarity: 83.33% Mismatches:  1
Query Match:    96.43%      Indels:      0
DB:             6           Gaps:        0

US-10-030-194A-5 (1-6) x CQ645844 (1-1239)

Qy      1  GlyTyr***ValGluGlu 6
      ||||| ||||| ||||| ||||| |||||
Db      916 GGTATAGTGTGAAGAA 933

RESULT 44
LOCUS       AX607019             1242 bp    DNA        linear    PAT 17-FEB-2003
DEFINITION   Sequence 4948 from Patent WO02092818.
ACCESSION    AX607019
VERSION      AX607019.1 GI:28402548
KEYWORDS     Streptococcus agalactiae
SOURCE       Streptococcus agalactiae
ORGANISM     Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE    1
AUTHORS      Glaser, P., Rusniok, C., Chevalier, F., Frangeul, L., Lalioui, L.,
Zouine, M., Couve, E., Buchrieser, C., Poyart, C., Trieu-Cuot, P. and
Kunst, F.
TITLE        Streptococcus agalactiae genome sequence, use for developing
vaccines, diagnostic tools, and for identifying therapeutic targets
JOURNAL      Patent: WO 02092818-A 4948 21-NOV-2002;
INSTITUT PASTEUR (FR) ; CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (FR)
FEATURES             Location/Qualifiers
     source          1..1242
                     /organism="Streptococcus agalactiae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1311"
ORIGIN
Alignment Scores:
Pred. No.:      409          Length:      1242
Score:          27.00        Matches:      5
Percent Similarity: 83.33%   Conservative: 0
Best Local Similarity: 83.33% Mismatches:  1
Query Match:    96.43%      Indels:      0
DB:             6           Gaps:        0

US-10-030-194A-5 (1-6) x AX607019 (1-1242)

Qy      1  GlyTyr***ValGluGlu 6
      ||||| ||||| ||||| ||||| |||||
Db      916 GGTATAGTGTGAAGAA 933

RESULT 45
LOCUS       CQ821806/c             1296 bp    DNA        linear    PAT 21-JUN-2004
DEFINITION   Sequence 151 from Patent WO2004048606.
ACCESSION    CQ821806
VERSION      CQ821806.1 GI:49020238
KEYWORDS     Sus scrofa (pig)
SOURCE       Sus scrofa
ORGANISM     Sus scrofa
REFERENCE    1
AUTHORS      Fukuyota; Metaxoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

```

```
REFERENCE 1
AUTHORS j Rgensen,C.B., Cirera,S., Archibald,A., Andersson,L., Fredholm,M.
TITLE Porcine polymorphisms and methods for detecting them
JOURNAL Patent: WO 2004048606-A 151 10-JUN-2004;
FEATURES Den KGL, Veterinaer- OG Landbohjskole (DK)
source Location/Qualifiers
1. 1296
/organism="Sus scrofa"
/mol_type="unassigned DNA"
/db_xref="taxon:9823"

ORIGIN
Alignment Scores:
Pred. No.: 430 Length: 1296
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x CQ821806 (1-1296)

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 900 GGATACTCAGTAGAAGAG 883

RESULT 46
SAU19899 1300 bp DNA linear BCT 25-MAR-1997
LOCUS Streptococcus agalactiae CMP-N-acetylneuraminic acid synthetase
DEFINITION (cpsF) gene, complete cds.
ACCESSION U19899
VERSION U19899.1 GI:642566
KEYWORDS Streptococcus agalactiae
SOURCE Streptococcus agalactiae
ORGANISM Streptococcus agalactiae
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 1300)
Hatt,R.F., Wessels,M.R., Mebane,M.F., Conaty,N. and Rubens,C.E.
AUTHORS Characterization of cpsF and its product CMP-N-acetylneuraminic
TITLE acid synthetase, a group B streptococcal enzyme that can function
in KI capsular polysaccharide biosynthesis in Escherichia coli
JOURNAL Mol. Microbiol. 19 (3), 555-563 (1996)
MEDLINE 96228704
PUBMED 8830246
REFERENCE 2 (bases 1 to 1300)
Rubens,C.E.
AUTHORS Direct Submission
TITLE Submitted (16-JAN-1995) Craig E. Rubens, Infectious Disease,
JOURNAL Children's Hospital and Medical Center, 4800 Sandpoint Way NE,
Seattle, WA 98105, USA
FEATURES Location/Qualifiers
source 1. .1300
/organism="Streptococcus agalactiae"
/mol_type="genomic DNA"
/strain="COH 31"
/db_xref="taxon:1311"
42. .45
52. .1239
/gene="cpsF"
52. .1239
/gene="cpsF"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="CMP-N-acetylneuraminic acid synthetase"
/protein_id="AAB50271.1"
/db_xref="GI:642567"
KQDIFVSDSELYREICLERGISVVMKPELSLTQATSYDMLKDFLSYDNQBFVLL
QVTSPLRKSWHIKEAMEYSSHDVNVVVSFSEVKEHPGLFTTLLSDKGVAIDMVGADKG
```

```
YRRDQLQYPYPNGAIFISNKETYLRKESFFTSYAYQMAKEFSLDVTRDDFIHVI
GHLFFDVAIREKENKVFYKGYSRFLRREASKIILGDSKTIISILENHYNSQCGVTI
ATMLENLPNFLTANVTAFVSIQVNDLIITGVSVVEIFSNFQKLYSLAENKI KWRFTT
IATYLFRETVNADIEKINQWLTEFCYQNPILLDINRFLSKDGNLYHLTSDGLHFD
SRG"

ORIGIN
Alignment Scores:
Pred. No.: 431 Length: 1300
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 1 Gaps: 0

US-10-030-194A-5 (1-6) x SAU19899 (1-1300)

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 967 GGTATATAGTGTGAAGAA 984

RESULT 47
S81109
LOCUS citrate synthase 1301 bp DNA linear BCT 14-MAR-2000
DEFINITION citrate synthase [Pyrococcus furiosus, Genomic, 1301 nt].
ACCESSION S81109
VERSION S81109.1 GI:1305725
KEYWORDS Pyrococcus furiosus
SOURCE Pyrococcus furiosus
ORGANISM Pyrococcus furiosus
Archaea; Euryarchaeota; Thermococci; Thermococcales;
Thermococcaceae; Pyrococcus.
REFERENCE 1 (bases 1 to 1301)
Muir,J.M., Russell,R.J., Hough,D.W. and Danson,M.J.
AUTHORS Citrate synthase from the hyperthermophilic Archaeon, Pyrococcus
TITLE furiosus
JOURNAL Protein Eng. 8 (6), 583-592 (1995)
MEDLINE 96081442
PUBMED 8532683
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibseq 175945] from the original journal article.
FEATURES Location/Qualifiers
source 1. .1301
/organism="Pyrococcus furiosus"
/mol_type="genomic DNA"
/db_xref="taxon:2261"
96. .1219
/gene="citrate synthase"
86. .1219
/gene="citrate synthase"
/note="Method: conceptual translation with partial peptide
sequencing; conceptual translation presented here differs
from translation in publication"
/codon_start=1
/transl_table=11
/product="citrate synthase"
/protein_id="AAB35835.2"
/db_xref="GI:7240619"
/transl_table=11
STFEVYLLWMGKLPSELSELENPKELAKSRGPKVEIEMALPKNTHPMGLARTI
ISYLGNDSDSDIGIPVTEPVYRIGISVTAKIPTIVANNYRIKNGLVYPPKELSHAA
NFLYMHGESPPKWEKAMDVALIYLAHEINASTLAVMTVGSTLSYSAIAGIGA
LKGP1HGGAVEEA1KQFMEIGSPKVEWFFKQQRKIMGAGRHYKTYDPRARIF
KKYASKLGDKKLFEIERLERLERLVEYLSSKGISINVDYWSGLVFGYGMKPIELVTTIF
AMGRIAGTAAHAEYVSHNRIIRPRLQYVQVGEIKKYLPIELRR"

ORIGIN
Alignment Scores:
Pred. No.: 432 Length: 1301
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43%
```

```

DB:
US-10-030-194A-5 (1-6) x S81109 (1-1301)
Gaps: 0

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 188 GGATACAGTGTAGAAGAG 205

RESULT 48
TC16SITAL/c
LOCUS TC16SITAL 1366 bp DNA linear BCT 02-DEC-1997
DEFINITION T. cecchii 16S rRNA gene, Italian strain.
ACCESSION Y09609
VERSION Y09609.1 GI:2661734
KEYWORDS 16S ribosomal RNA; 16S rRNA.
SOURCE Tetracoccus cecchii
ORGANISM Tetracoccus cecchii
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
Rhodobacteraceae; Tetracoccus.
REFERENCE 1
AUTHORS Blackall,L.L., Rossetti,S., Christenson,C., Cunningham,M.,
Hartman,P., Hugenholtz,P. and Tandoi,V.
TITLE The characterization and description of representatives of 'G'
bacteria from activated sludge plants
JOURNAL Lett. Appl. Microbiol. 25 (1), 63-69 (1997)
MEDLINE 97391276
PUBMED 9248083
REFERENCE 2 (bases 1 to 1366)
AUTHORS Blackall,L.L.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-1996) L.L. Blackall, The University of
Queensland, Department of Microbiology, Brisbane, 4072, Queensland,
AUSTRALIA

FEATURES
source Location/Qualifiers
1..1366
/mol_type="genomic DNA"
/strain="Italian"
/db_xref="taxon:56057"
1..1366
/gene="16S rRNA"
<1..>1366
/gene="16S rRNA"
/product="16S ribosomal RNA"

ORIGIN
Alignment Scores: 456 Length: 1366
Pred. No.: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 1 Gaps: 0

US-10-030-194A-5 (1-6) x TC16SITAL (1-1366)
Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 121 GGCTATTCGTAGAAGAG 104

RESULT 49
AX646063/c
LOCUS AX646063 1387 bp DNA linear PAT 04-MAR-2003
DEFINITION Sequence 255 from Patent EP1270724.
ACCESSION AX646063
VERSION AX646063.1 GI:28798442
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Suwa,M., Asai,K., Akiyama,Y. and Aburatani,H.

```

```

TITLE Guanosine triphosphate-binding protein coupled receptors
JOURNAL Patent: EP 1270724-A 255 02-JAN-2003;
National Institute of Advanced Industrial Science and Technology
(JP) ; Center for Advanced Science and Technology Incubation, Ltd.
(JP)

FEATURES
source Location/Qualifiers
1..1387
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
201..1187
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD69167.1"
/db_xref="GI:28798443"
/translation="MPLTERNTTSEATFTLLIGSDYLELQIPLPFVFLAVYGFVVGNG
LGMVIIKINPKLHTPMYFFLNHLSEVDFCVSSIIAPMMLVNLVVEDRTISFGCLVQ
FFPCTFVTELLILFAMAYDHFVAICNPLLYTVAISQKLCAMLVVLVYANGVACSLT
LACALKSLDFHGFNTINHFCELSLSLSPDSYLSQLLEFTVATFNEISTLLILIT
SYAFIIVTTLKMPASGHRKVFTSCASHLTAITIFHGTLILELYCVFNSKSRHTVKVA
SVFYTVVIVPLNPLIYSLRNKNDKDAIRKIINTKYFHKHRHWPFPNFVIEQ"

ORIGIN
Alignment Scores: 464 Length: 1387
Pred. No.: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x AX646063 (1-1387)
Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 992 GGTACACAGTAGAGAA 975

RESULT 50
AB065783/c
LOCUS AB065783 1387 bp DNA linear PRI 23-JUL-2002
DEFINITION Homo sapiens gene for seven transmembrane helix receptor, complete
cds, isolate:CBRC7TM_346.
ACCESSION AB065783
VERSION AB065783.1 GI:21928833
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S.,
Tateumi,S., Aburatani,H., Asai,K. and Akiyama,Y.
TITLE Genome-wide discovery and analysis of human seven transmembrane
helix receptor genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1387)
AUTHORS Suwa,M.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research
Center (CBRC), National Institute of Advanced Industrial Science
and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan
(E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/,
Tel:81-3-3599-8080, Fax:81-3-3599-8081)
COMMENT This sequence is a seven transmembrane helix receptor candidate
predicted from the whole human genome sequences using our automated
system that contains programs of gene
finding(GeneDecoder), sequence search, motif-domain assignment and
transmembrane helix prediction.
And the sequence is submitted by the collaborative project between
[Computational Biology Research Center (CBRC), National Institute
of Advanced Industrial Science and Technology (AIST)] and [Genome
Science Division, Research Center for Advanced Science and
Technology (RCAST), University of Tokyo].

```

```

FEATURES
  source      Location/Qualifiers
              1..1387
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /isolate="CBRC7TM 346"
              /db_xref="taxon:9606"
              /chromosome="11"
              201..1187
              /codon_start=1
              /evidence=not_experimental
              /product="seven transmembrane helix receptor"
              /protein_id="BAC06003.1"
              /db_xref="GI:21928834"
              /translation="MELTERTNTSEATFTLLGFSDDYLELOIPLFVFLAVVGFSGVGN
              LGMIVIKIPKHTPMYFELNHLSEFVDFCYSSIIAPMLVNLVVEDRTISFGCLVQ
              FFFCTFVTEILIFAWADHFVAICNPLIYVAISQKLCAMLVVLYANGVACSIT
              LACSALKSLFHGFTINHHFCELSLSLSYPDSYLSQLLFTVATNEISTLLIIT
              SYAFIVTLKMPASGHRKVFSTCASHLTATITFHGTLIFLFCVPSKNSRHTKVA
              SVFYTVIPLNPLIYSLRNKVDKDAIRKIINTKYFHKRHMYPPNFVIEQ"

CDS
  Alignment Scores:
  Pred. No.:      464      Length:      1387
  Score:          27.00     Matches:      5
  Percent Similarity: 83.33%  Conservative: 0
  Best Local Similarity: 83.33%  Mismatches: 1
  Query Match:    96.43%      Indels:      0
  DB:             9        Gaps:          0

ORIGIN
  US-10-030-194A-5 (1-6) x AB065783 (1-1387)
  Qy      1 GlyTyr***ValGluGlu 6
  Db      992 GGGTACACAGTAGAGGAA 975

RESULT 51
AY034379
LOCUS      AY034379      1398 bp      mRNA      linear      PLN 03-JUN-2001
DEFINITION Capsicum annuum branched-chain amino acid aminotransferase mRNA,
complete cds.
ACCESSION      AY034379
VERSION      AY034379.1 GI:14280353
KEYWORDS
SOURCE      Capsicum annuum
ORGANISM      Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Capsicum.
REFERENCE
  1 (bases 1 to 1398)
  Balducci,E., Richins,R.D., Curry,J. and O'Connell,M.A.
  Nucleotide sequence of a branched-chain amino acid aminotransferase
  from chile challenged with Phytophthora capsici
  JOURNAL      Unpublished
REFERENCE
  2 (bases 1 to 1398)
  Balducci,E., Richins,R.D., Curry,J. and O'Connell,M.A.
  Direct Submission
  TITLE      Direct Submission
  JOURNAL      Submitted (11-MAY-2001) Agronomy & Horticulture, New Mexico State
  University, MSC 3Q, P.O. Box 30003, Las Cruces, NM 88003, USA
FEATURES
  source      Location/Qualifiers
              1..1398
              /organism="Capsicum annuum"
              /mol_type="mRNA"
              /cultivar="CM334"
              /db_xref="taxon:4072"
              85..1242
              /note="BCAT"
              /codon_start=1
              /product="branched-chain amino acid aminotransferase"
              /protein_id="AAK57535.1"
              /db_xref="GI:14280354"
              /translation="MIRGAACFRKLFOSSALSSKFTTRYTAAQVAPAIYSSDEESG
              SIDWNLGFKLMQTDYMYTKTSDGI FRQGLDRYGNIQLSPSAGVNLGYQGLFEGT
              KAFREDGRFLFRPEQNAIRMQIGAERMCMFAPSTDQFVDALKQTALSNKRWTPPPG"

CDS
  Alignment Scores:
  Pred. NO.:      477      Length:      1421
  Score:          27.00     Matches:      5
  Percent Similarity: 83.33%  Conservative: 0
  Best Local Similarity: 83.33%  Mismatches: 1
  Query Match:    96.43%      Indels:      0
  DB:             1        Gaps:          0

US-10-030-194A-5 (1-6) x TC16SCZEC (1-1421)
  Qy      1 GlyTyr***ValGluGlu 6
  Db      124 GGCTATTCCGTAGAAGAG 107

RESULT 52
TC16SCZEC/c
LOCUS      TC16SCZEC      1421 bp      DNA      linear      BCT 02-DEC-1997
DEFINITION T.cechii 16S rRNA gene, Czech strain.
ACCESSION      Y09610
VERSION      Y09610.1 GI:2661735
KEYWORDS      16S ribosomal RNA; 16S rRNA.
SOURCE      Tetracoccus cechii
ORGANISM      Tetracoccus cechii
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
Rhodobacteraceae; Tetracoccus.
REFERENCE
  1
  AUTHORS      Blackall,L.L., Rossetti,S., Christensson,C., Cunningham,M.,
  Hartman,P., Hugenholtz,P. and Tandoi,V.
  TITLE      The characterization and description of representatives of 'G'
  bacteria from activated sludge plants
  JOURNAL      Lett. Appl. Microbiol. 25 (1), 63-69 (1997)
  MEDLINE      97391276
  PUBMED      9248083
REFERENCE
  2 (bases 1 to 1421)
  Blackall,L.L.
  Direct Submission
  TITLE      Submitted (25-NOV-1996) L.L. Blackall, The University of
  Queensland, Department of Microbiology, Brisbane, 4072, Queensland,
  AUSTRALIA
FEATURES
  source      Location/Qualifiers
              1..1421
              /organism="Tetracoccus cechii"
              /mol_type="genomic DNA"
              /strain="Czech"
              /db_xref="taxon:56057"
              1..1421
              /gene="16S rRNA"
              <1..>1421
              /gene="16S rRNA"
              /product="16S ribosomal RNA"

ORIGIN
  Alignment Scores:
  Pred. NO.:      477      Length:      1421
  Score:          27.00     Matches:      5
  Percent Similarity: 83.33%  Conservative: 0
  Best Local Similarity: 83.33%  Mismatches: 1
  Query Match:    96.43%      Indels:      0
  DB:             1        Gaps:          0

US-10-030-194A-5 (1-6) x AY034379 (1-1398)
  Qy      1 GlyTyr***ValGluGlu 6
  Db      985 GGATATACAGGTTGAAGAA 1002

```

```

KGSLYIRPLLIGTGPLGLAPAPETFLVYACPVGNVYFKEGTAPFLNLYVEDVHRASR
GGAGGVKSIITNYPVLKAMKXKANGSDLYVDVANKYIIEVSSCNIFVYKGNVVS
TPIAKGTILGVTTRKSTMDIALDGYTVBERLIIEADELVCTGTAVGAPVGS
ITYKGRIEYKISSDLSCKFPYSRLVGIQGVIKDERNWIVEIESSYLVD"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      468      Length:      1398
Score:          27.00     Matches:      5
Percent Similarity: 83.33%  Conservative: 0
Best Local Similarity: 83.33%  Mismatches: 1
Query Match:    96.43%      Indels:      0
DB:             8        Gaps:          0

```

US-10-030-194A-5 (1-6) x AY034379 (1-1398)

```

Qy      1 GlyTyr***ValGluGlu 6
      ||||| ||||| ||||| |||||
Db      985 GGATATACAGGTTGAAGAA 1002

```

RESULT 52

```

TC16SCZEC/c
LOCUS      TC16SCZEC      1421 bp      DNA      linear      BCT 02-DEC-1997
DEFINITION T.cechii 16S rRNA gene, Czech strain.
ACCESSION      Y09610
VERSION      Y09610.1 GI:2661735
KEYWORDS      16S ribosomal RNA; 16S rRNA.
SOURCE      Tetracoccus cechii
ORGANISM      Tetracoccus cechii
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
Rhodobacteraceae; Tetracoccus.
REFERENCE
  1
  AUTHORS      Blackall,L.L., Rossetti,S., Christensson,C., Cunningham,M.,
  Hartman,P., Hugenholtz,P. and Tandoi,V.
  TITLE      The characterization and description of representatives of 'G'
  bacteria from activated sludge plants
  JOURNAL      Lett. Appl. Microbiol. 25 (1), 63-69 (1997)
  MEDLINE      97391276
  PUBMED      9248083
REFERENCE
  2 (bases 1 to 1421)
  Blackall,L.L.
  Direct Submission
  TITLE      Submitted (25-NOV-1996) L.L. Blackall, The University of
  Queensland, Department of Microbiology, Brisbane, 4072, Queensland,
  AUSTRALIA
FEATURES
  source      Location/Qualifiers
              1..1421
              /organism="Tetracoccus cechii"
              /mol_type="genomic DNA"
              /strain="Czech"
              /db_xref="taxon:56057"
              1..1421
              /gene="16S rRNA"
              <1..>1421
              /gene="16S rRNA"
              /product="16S ribosomal RNA"

```

ORIGIN

```

Alignment Scores:
Pred. NO.:      477      Length:      1421
Score:          27.00     Matches:      5
Percent Similarity: 83.33%  Conservative: 0
Best Local Similarity: 83.33%  Mismatches: 1
Query Match:    96.43%      Indels:      0
DB:             1        Gaps:          0

US-10-030-194A-5 (1-6) x TC16SCZEC (1-1421)

```

```

Qy      1 GlyTyr***ValGluGlu 6
      ||||| ||||| ||||| |||||
Db      124 GGCTATTCCGTAGAAGAG 107

```

RESULT 53

AKU88041/c	AKU88041	1428 bp	DNA	linear	BCT 21-AUG-1997
LOCUS	Amariococcus kaplicensis	16S ribosomal RNA gene, complete sequence.			
DEFINITION	U88041				
ACCESSION	U88041.1	GI:2340153			
VERSION					
KEYWORDS	Amariococcus kaplicensis				
SOURCE	Amariococcus kaplicensis				
ORGANISM	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Amariococcus.				
REFERENCE	1 (bases 1 to 1428)				
AUTHORS	Mazenan,A.M., Seviour,R.J., Patel,B.K., Rees,G.N. and McDougall,B.M.				
TITLE	Amariococcus gen. nov., a gram-negative coccus occurring in regular packages or tetrads, isolated from activated sludge biomass, and descriptions of Amariococcus veronensis sp. nov., Amariococcus tamworthensis sp. nov., Amariococcus macauensis sp. nov., and Amariococcus kaplicensis sp. nov				
JOURNAL	Int. J. Syst. Bacteriol. 47 (3), 727-734 (1997)				
MEDLINE	97370598				
PUBMED	9226904				
REFERENCE	2 (bases 1 to 1428)				
AUTHORS	Patel,B.K.C. and Mazenan,A.M.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-FEB-1997) School of Biomolecular and Biomedical Sciences, Griffith University, Nathan, Brisbane, Queensland 4111, Australia				
FEATURES	Location/Qualifiers				
source	1..1428				
	/organism="Amariococcus kaplicensis"				
	/mol_type="genomic DNA"				
	/strain="Ben101"				
	/db_xref="taxon:57000"				
rRNA	1..1428				
	/product="16S ribosomal RNA"				
ORIGIN					
Alignment Scores:					
Pred. No.:	480	Length:	1428		
Score:	27.00	Matches:	5		
Percent Similarity:	83.33%	Conservative:	0		
Best Local Similarity:	83.33%	Mismatches:	1		
Query Match:	96.43%	Indels:	0		
DB:	1	Gaps:	0		
US-10-030-194A-5 (1-6) x AKU88041 (1-1428)					
Qy	1 GlyTyr***ValGluGlu 6				
Db	128 GGCATTCCGTAGAGAG 111				
RESULT 54					
FSU23722					
LOCUS	Fusarium solani pisi cutinase gene palindrome-binding protein mRNA, complete cds.	1430 bp	mRNA	linear	PLN 26-JAN-1996
DEFINITION					
ACCESSION	U23722				
VERSION	U23722.1	GI:763041			
KEYWORDS	cutinase gene palindrome-binding protein.				
SOURCE	Nectria haematococca				
ORGANISM	Nectria haematococca				
REFERENCE	1 (bases 1 to 1430)				
AUTHORS	Li,D. and Kolattukudy,P.E.				
TITLE	Cloning and expression of cDNA encoding a protein that binds a palindromic promoter element essential for induction of fungal cutinase by plant cutin				
JOURNAL	J. Biol. Chem. 270 (20), 11753-11756 (1995)				
MEDLINE	95263512				
PUBMED	7744822				
REFERENCE	2 (bases 1 to 1430)				
AUTHORS	Li,D.				
Alignment Scores:					
Pred. No.:	498	Length:	1476		
Score:	27.00	Matches:	5		
Percent Similarity:	83.33%	Conservative:	0		
Best Local Similarity:	83.33%	Mismatches:	1		
Query Match:	96.43%	Indels:	0		
DB:	8	Gaps:	0		
US-10-030-194A-5 (1-6) x FSU23722 (1-1430)					
Qy	1 GlyTyr***ValGluGlu 6				
Db	451 GGGTACAGCGTGGAGGAG 468				
RESULT 55					
CQ729805/c					
LOCUS	CQ729805	1476 bp	DNA	linear	PAT 03-FEB-2004
DEFINITION	Sequence 15739 from Patent WO02068579.				
ACCESSION	CQ729805				
VERSION	CQ729805.1	GI:42301899			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.				
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof				
JOURNAL	Patent: WO 02068579-A 15739 06-SEP-2002;				
PE Corporation (NY) (US)					
FEATURES	Location/Qualifiers				
source	1..1476				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Alignment Scores:					
Pred. No.:	498	Length:	1476		

Direct Submission					
Submitted (29-MAR-1995)	Daoxin Li, Neurobiotechnology, The Ohio State University, 206 Rightmire Hall, 1060 Carmack Rd, Columbus, OH 43210, USA				
FEATURES	Location/Qualifiers				
source	1..1430				
	/organism="Nectria haematococca"				
	/mol_type="mRNA"				
	/strain="T8"				
	/db_xref="taxon:140110"				
	/clone="ppbp"				
	/clone_lib="lamda gt-11"				
	37..1410				
CDS					
	/codon_start=1				
	/product="cutinase gene palindrome-binding protein"				
	/protein_id="AAA85727.1"				
	/db_xref="GI:763042"				
	/translation="MTGGQMGADGQSDTDIMALLDNNMLGSGDVSMGLDAVDESSMT				
	GSFGSNAQTSTESVARAENQFNTAANNVPGAGLPAGPFAQNNLTSAASLTLEFTK				
	RNNWPAKVVEELRDFLQILDANGRIKYAPSPILNVTGYSVEEIQDILKDLIHPDDR				
	VFVAELNESIASGNPLRLFYFKKDGKFAIFETVGHAIAGSKFAPNPNQSPFQCA				
	VFMARPYPTKNAGLLDSFLEHKIENERLKRRIAEALREBEAENDEAKQWQSQEGR				
	SDVTPSEGTGVSSNFFRPSTDRGLTDADRALNKSLLTRENLEGAAGSRPDSLNDKMAR				
	YEGASHTETIEMLTGLRYIEGERSRGITTCGNASPTLTKGDAGIALPLERDPRTDGKKK				
	KLKTSSEYVCTDCGTLDSPWVRKGPSPKTLCLNACGLRWAKKCKRNSSVNHRAPGP				
	AIDH5MA"				
misc_feature	1200..1221				
	/note="encodes potential nuclear localization signal"				
misc_feature	1242..1323				
	/note="encodes 1 GATA-like zinc finger motif"				
ORIGIN					
Alignment Scores:					
Pred. No.:	480	Length:	1430		
Score:	27.00	Matches:	5		
Percent Similarity:	83.33%	Conservative:	0		
Best Local Similarity:	83.33%	Mismatches:	1		
Query Match:	96.43%	Indels:	0		
DB:	8	Gaps:	0		
US-10-030-194A-5 (1-6) x FSU23722 (1-1430)					
Qy	1 GlyTyr***ValGluGlu 6				
Db	451 GGGTACAGCGTGGAGGAG 468				
RESULT 55					
CQ729805/c					
LOCUS	CQ729805	1476 bp	DNA	linear	PAT 03-FEB-2004
DEFINITION	Sequence 15739 from Patent WO02068579.				
ACCESSION	CQ729805				
VERSION	CQ729805.1	GI:42301899			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.				
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof				
JOURNAL	Patent: WO 02068579-A 15739 06-SEP-2002;				
PE Corporation (NY) (US)					
FEATURES	Location/Qualifiers				
source	1..1476				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Alignment Scores:					
Pred. No.:	498	Length:	1476		

```
Score:                27.00           Matches:                5
Percent Similarity:   83.33%          Conservative:         0
Best Local Similarity: 83.33%          Mismatches:          1
Query Match:          96.43%          Indels:              0
DB:                   6               Gaps:               0

US-10-030-194A-5 (1-6) x CQ729805 (1-1476)

QY      1  GlyTyr***ValGluGlu 6
Db      1281 GGTACACAGTAGAGGAA 1264

RESULT 56
MAUCC10 MAUCC10          1491 bp      DNA      linear      ROD 23-MAY-1998
LOCUS   Mesocricetus auratus ug/cc10 gene.
DEFINITION Y13765
ACCESSION Y13765
VERSION   Y13765.1 GI:3157403
KEYWORDS  ug/cc10 gene; uteroglobin/clara cell 10kDa protein.
SOURCE    Mesocricetus auratus (golden hamster)
ORGANISM  Mesocricetus auratus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
          Mesocricetus.
REFERENCE 1
AUTHORS   Sagal,R.G. and Nieto,A.
TITLE     Molecular cloning of the cDNA and the promoter of the hamster
          uteroglobin/clara cell 10-kDa gene (ug/cc10): tissue-specific and
          hormonal regulation
JOURNAL   Arch. Biochem. Biophys. 350 (2), 214-222 (1998)
MEDLINE   98141958
PUBMED   9473294
REFERENCE 2 (bases 1 to 1491)
AUTHORS   Gutierrez-Sagal,R.
TITLE     Direct Submission
JOURNAL   Submitted (11-JUN-1997) R. Gutierrez-Sagal, Centro de Biologia
          Molecular, Universidad Autonoma, Canto Blaco, 28049 Madrid, SPAIN
COMMENT   Related entry: I37041.
FEATURES  Location/Qualifiers
          source             1..1491
                           /organism="Mesocricetus auratus"
                           /mol_type="genomic DNA"
                           /db_xref="taxon:10036"
                           /dev_stage="adult"
          gene               1363..1491
                           /gene="ug/cc10"
          TATA_signal        1363..1366
                           /gene="ug/cc10"
          mRNA               1393..>1491
                           /gene="ug/cc10"
          prim_transcript    1393..>1491
                           /gene="ug/cc10"
          CDS                1450..>1491
                           /gene="ug/cc10"
                           /codon_start=1
                           /product="uteroglobin/clara cell 10kDa protein"
                           /protein_id="CAA74098.1"
                           /db_xref="GI:3157404"
                           /db_xref="TrEMBL:O70544"
                           /translation="MKIAITMAVVMLSV"

ORIGIN
Alignment Scores:
Pred. No.:          504           Length:          1491
Score:              27.00         Matches:          5
Percent Similarity: 83.33%        Conservative:    0
Best Local Similarity: 83.33%      Mismatches:     1
Query Match:        96.43%        Indels:         0
DB:                 10           Gaps:           0

US-10-030-194A-5 (1-6) x MAUCC10 (1-1491)
QY      1  GlyTyr***ValGluGlu 6

Score:                27.00           Matches:                5
Percent Similarity:   83.33%          Conservative:         0
Best Local Similarity: 83.33%          Mismatches:          1
Query Match:          96.43%          Indels:              0
DB:                   6               Gaps:               0

US-10-030-194A-5 (1-6) x AX078547 (1-1497)
QY      1  GlyTyr***ValGluGlu 6
Db      16 GGCTACAGTGTGGAAGAG 33

RESULT 58
XELTRANSD
LOCUS   Xenopus laevis (clone pXgal) transducin alpha subunit mRNA,
DEFINITION complete cds.
ACCESSION L07771
VERSION   L07771.1 GI:214837
KEYWORDS  G protein; guanylate binding protein; phototransduction protein;
          retinal protein; transducin alpha subunit.
          Xenopus laevis (African clawed frog)
SOURCE    Xenopus laevis
ORGANISM  Xenopus laevis
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
          Xenopodinae; Xenopus; Xenopus.
          1 (bases 1 to 1507)
          Knox,B.E., Scalzetti,L.C., Batni,S. and Wang,J.Q.
          Molecular cloning of the abundant rhodopsin and transducin from
          Xenopus laevis
          Unpublished (1992)
          Original source text: Xenopus laevis (tissue library: lambda-ZAPII)
          adult retina cDNA to mRNA.
          Location/Qualifiers
          source             1..1507
                           /organism="Xenopus laevis"
                           /mol_type="mRNA"
                           /db_xref="taxon:8355"
                           /tissue_type="retina"

US-10-030-194A-5 (1-6) x AX078547 (1-1497)
QY      1  GlyTyr***ValGluGlu 6
Db      16 GGCTACTCTGTGGAGGAA 37

RESULT 57
XELTRANSD
LOCUS   Xenopus laevis (clone pXgal) transducin alpha subunit mRNA,
DEFINITION complete cds.
ACCESSION L07771
VERSION   L07771.1 GI:214837
KEYWORDS  G protein; guanylate binding protein; phototransduction protein;
          retinal protein; transducin alpha subunit.
          Xenopus laevis (African clawed frog)
SOURCE    Xenopus laevis
ORGANISM  Xenopus laevis
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
          Xenopodinae; Xenopus; Xenopus.
          1 (bases 1 to 1507)
          Fraser,C.M., Venter,C., Tuemmler,B., Hoheisel,J., Duesterhoeft,A.,
          Hilbert,H., Timmis,K.N., Moore,E., Straetz,M., Heim,S.,
          Nelson,K.E., Hickey,E. and Peterson,J.
          Dna sequences which are suited for specifically detecting
          Pseudomonas putida kt2440
          Patent: WO 0107624-A 61 01-FEB-2001;
          THE INSTITUTE FOR GENOMIC RESEARCH (US) ; QIAGEN GmbH (DE) ;
          Gesellschaft fuer Biotechnologische Forschung mbH (GBF) (DE) ;
          Deutsches Krebsforschungszentrum (DKFZ) (DE) ; Medizinische
          Hochschule Hannover (DE)
          Location/Qualifiers
          source             1..1497
                           /organism="Pseudomonas putida"
                           /mol_type="unassigned DNA"
                           /db_xref="taxon:303"

ORIGIN
Alignment Scores:
Pred. No.:          506           Length:          1497
Score:              27.00         Matches:          5
Percent Similarity: 83.33%        Conservative:    0
Best Local Similarity: 83.33%      Mismatches:     1
Query Match:        96.43%        Indels:         0
DB:                 6               Gaps:           0

US-10-030-194A-5 (1-6) x AX078547 (1-1497)
QY      1  GlyTyr***ValGluGlu 6
Db      16 GGCTACTCTGTGGAGGAA 37

RESULT 57
XELTRANSD
LOCUS   Xenopus laevis (clone pXgal) transducin alpha subunit mRNA,
DEFINITION complete cds.
ACCESSION L07771
VERSION   L07771.1 GI:214837
KEYWORDS  G protein; guanylate binding protein; phototransduction protein;
          retinal protein; transducin alpha subunit.
          Xenopus laevis (African clawed frog)
SOURCE    Xenopus laevis
ORGANISM  Xenopus laevis
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
          Xenopodinae; Xenopus; Xenopus.
          1 (bases 1 to 1507)
          Knox,B.E., Scalzetti,L.C., Batni,S. and Wang,J.Q.
          Molecular cloning of the abundant rhodopsin and transducin from
          Xenopus laevis
          Unpublished (1992)
          Original source text: Xenopus laevis (tissue library: lambda-ZAPII)
          adult retina cDNA to mRNA.
          Location/Qualifiers
          source             1..1507
                           /organism="Xenopus laevis"
                           /mol_type="mRNA"
                           /db_xref="taxon:8355"
                           /tissue_type="retina"

US-10-030-194A-5 (1-6) x MAUCC10 (1-1491)
QY      1  GlyTyr***ValGluGlu 6
```

```
/dev_stage="adult"
/tissue_lib="lambda-ZAPII"
150..1202
/codon_start=1
/product="transducin alpha subunit"
/protein_id="AAA88693.1"
/db_xref="GI:214838"
/translation="MGASAEKHSRELEKKBADKADARTVKLLLLGAGESGKST
IVQKMTIHODGYSVEEELFTAIYIGNTLQSMIAIVKAMTNLQFGDPARQDDARK
LMLADITDEGSMPEKMDIIGRLWKDTGQACFRASEYQLNDSAGYLLDLRLVI
PGYPTEDVLRSRVKTGIIETOPGLKDLNRFMDVGQSRKWTIHCFEGVTCII
PIALSAYDVLVDEVENHMSHLFNSICNHRYPATTSIVLFLNKKDVPTEKIKK
AHLSTCFPDYDGPNTYDAGNYIKTQFLNLRDRDKSIYSHMTCAIDTENVKVFDA
VTDIIKENLKDGLF"
misc_feature
325..1507
/note="sequenced from clone pXgal"

ORIGIN
Alignment Scores:
Pred. No.: 510 Length: 1507
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x XELTRANS (1-1507)

QY 1 GlyTyr***ValGluGlu 6
DB 315 GGGTACTCGTTGAGGAA 332

RESULT 59
SHU12238 1529 bp DNA linear BCT 31-JUL-1994
LOCUS Staphylococcus haemolyticus Y176 D-amino acid transaminase (dat)
DEFINITION gene, complete cds.
ACCESSION U12238
VERSION U12238.1 GI:517474
KEYWORDS
SOURCE Staphylococcus haemolyticus
ORGANISM Staphylococcus haemolyticus
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1 (bases 1 to 1529)
AUTHORS Pucci,M.J., Thanassi,J.A., Ho,H., Falk,P.J. and Dougherty,T.J.
TITLE Staphylococcus haemolyticus contains two D-glutamic acid
biosynthetic activities: a glutamate racemase and a D-amino acid
transaminase
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1529)
AUTHORS Pucci,M.J.
TITLE Direct Submission
JOURNAL Submitted (12-JUL-1994) Michael J. Pucci, Microbiology,
Bristol-Myers Squibb Company, 5 Research Pkwy., Wallingford, CT
06492-7660, USA

FEATURES
source
1..1529
Location/Qualifiers
/organism="Staphylococcus haemolyticus"
/mol_type="genomic DNA"
/strain="Y176"
/db_xref="taxon:1283"
204..209
216..1064
/gene="dat"
216..1064
/gene="dat"
/codon_start=1
/transl_table=11
/product="D-amino acid transaminase"
/protein_id="AAA20396.1"
/db_xref="GI:517475"
/translation="MTKVFINGEFTDQNEAKVSYEDRGYVFGDGIYEIRAYDGKLF
VTEHFERIRGASEIQLDLGYTVBELIDVRELLKVNINQGGIYQATRGVAPRNHS

/dev_stage="adult"
/tissue_lib="lambda-ZAPII"
150..1202
/codon_start=1
/product="transducin alpha subunit"
/protein_id="AAA88693.1"
/db_xref="GI:214838"
/translation="MGASAEKHSRELEKKBADKADARTVKLLLLGAGESGKST
IVQKMTIHODGYSVEEELFTAIYIGNTLQSMIAIVKAMTNLQFGDPARQDDARK
LMLADITDEGSMPEKMDIIGRLWKDTGQACFRASEYQLNDSAGYLLDLRLVI
PGYPTEDVLRSRVKTGIIETOPGLKDLNRFMDVGQSRKWTIHCFEGVTCII
PIALSAYDVLVDEVENHMSHLFNSICNHRYPATTSIVLFLNKKDVPTEKIKK
AHLSTCFPDYDGPNTYDAGNYIKTQFLNLRDRDKSIYSHMTCAIDTENVKVFDA
VTDIIKENLKDGLF"
misc_feature
325..1507
/note="sequenced from clone pXgal"

ORIGIN
Alignment Scores:
Pred. No.: 510 Length: 1507
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x XELTRANS (1-1507)

QY 1 GlyTyr***ValGluGlu 6
DB 315 GGGTACTCGTTGAGGAA 332

RESULT 59
SHU12238 1529 bp DNA linear BCT 31-JUL-1994
LOCUS Staphylococcus haemolyticus Y176 D-amino acid transaminase (dat)
DEFINITION gene, complete cds.
ACCESSION U12238
VERSION U12238.1 GI:517474
KEYWORDS
SOURCE Staphylococcus haemolyticus
ORGANISM Staphylococcus haemolyticus
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1 (bases 1 to 1529)
AUTHORS Pucci,M.J., Thanassi,J.A., Ho,H., Falk,P.J. and Dougherty,T.J.
TITLE Staphylococcus haemolyticus contains two D-glutamic acid
biosynthetic activities: a glutamate racemase and a D-amino acid
transaminase
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1529)
AUTHORS Pucci,M.J.
TITLE Direct Submission
JOURNAL Submitted (12-JUL-1994) Michael J. Pucci, Microbiology,
Bristol-Myers Squibb Company, 5 Research Pkwy., Wallingford, CT
06492-7660, USA

FEATURES
source
1..1529
Location/Qualifiers
/organism="Staphylococcus haemolyticus"
/mol_type="genomic DNA"
/strain="Y176"
/db_xref="taxon:1283"
204..209
216..1064
/gene="dat"
216..1064
/gene="dat"
/codon_start=1
/transl_table=11
/product="D-amino acid transaminase"
/protein_id="AAA20396.1"
/db_xref="GI:517475"
/translation="MTKVFINGEFTDQNEAKVSYEDRGYVFGDGIYEIRAYDGKLF
VTEHFERIRGASEIQLDLGYTVBELIDVRELLKVNINQGGIYQATRGVAPRNHS

PPTDEVKPVINAFKASYDRPYDDLENGINATVEDIRWLRCDIKSLNLLGNVLAKKEYA
VKYNAGAIQHRGETVTEGASSNYIAIKDGAITYHPVNNYILNGITRKVIWISSEDED
IPFKETFTVEFLKNADEVIVSSTSAEVTVPVKIDGQVGDKGVPTROLQEGFNKY
IESRSS"

ORIGIN
Alignment Scores:
Pred. No.: 518 Length: 1529
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 1 Gaps: 0

US-10-030-194A-5 (1-6) x SHU12238 (1-1529)

QY 1 GlyTyr***ValGluGlu 6
DB 405 GGCTATACTGTCGAAGAA 422

RESULT 60
BC046071 1531 bp mRNA linear VRT 30-JUN-2004
LOCUS Danio rerio zgc:56319, mRNA (cdna clone MGC:56319 IMAGE:5603467),
DEFINITION complete cds.
ACCESSION BC046071
VERSION BC046071.1 GI:28374343
KEYWORDS MGC.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

REFERENCE 1 (bases 1 to 1531)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,J., Shenmen,C.M., Schuler,G.D.,
Altshuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toehiyuki,S.,
Carninci,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 1531)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Sumio Sugano
cDNA Library Preparation: Dr. Sumio Sugano
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
```

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 103 Row: 0 Column: 11

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES
source
1. .1531
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:56319 IMAGE:5603467"
/tissue_type="Whole body, adult male"
/clone_lib="Sugano SJD adult male"
/lab_host="DH10B"
/note="vector: pME18S-FL3"
gene
1. .1531
/gene="zgc:56319"
/note="synonym: MGCS6319"
/db_xref="LocusID:393198"
45. .620
/gene="zgc:56319"
/codon_start=1
/product="similar to COP9 homolog"
/protein_id="AAH46071.1"
/db_xref="GI:28374344"
/translation="MPVCVMMAELDKLLQFETQELAPGGIATPQVYSQLLVLYLL
HNDNNARYLWKRIPOAKITANPEMAAIWAGQRIWQDFPGIYSAIAAYQWSSILP
VMEALRESTRERRVGLVAQAVTSTSAEDFAAFVGSVEEAVKGVVSHGWQADPNTRMI
MPQKDPDPVSLVNEQQALRTDYVALEN"
ORIGIN

Alignment Scores:
Pred. No.: 519 Length: 1531
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BC046071 (1-1531)

Qy 1 GlyTyr***ValGluGlu 6
Db 450 GGATACTCGTAGAGAG 467

RESULT 61
SCCYTB5R
LOCUS SCCYTB5R 1586 bp DNA linear PLN 16-MAR-1994
DEFINITION S.cerevisiae (842) gene for cytochrome b5 reductase.
ACCESSION Z28365
VERSION Z28365.1 GI:461337
KEYWORDS cytochrome b5 reductase.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 1586)
AUTHORS Csukai,M., Murray,M. and Orr,E.
TITLE Isolation and complete sequence of CBR, a gene encoding a putative cytochrome b reductase in Saccharomyces cerevisiae
JOURNAL Eur. J. Biochem. 219 (1-2), 441-448 (1994)
MEDLINE 94139721
PUBMED 8307010
REFERENCE 2 (bases 1 to 1586)
AUTHORS Csukai,M.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-1993) Michael Csukai, Department of Genetics, University of Leicester, University Road, Leicester,

Leicestershire, LE1 7RH, U.K
Location/Qualifiers
1. .1586
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="842"
/db_xref="taxon:4932"
/chromosome="Chromosome IX"
/clone="pCBR"
/clone_lib="lambda gt11"
385. .1353
/gene="cytochrome b5 reductase"
385. .1353
/gene="cytochrome b5 reductase"
/codon_start=1
/protein_id="CAA82214.1"
/db_xref="GI:461338"
/db_xref="GOA:P38626"
/db_xref="SGD:S0001305"
/db_xref="Swiss-Prot:P38626"
/translation="MYKSYVIRKNERKKVLKVCIALQQLQQTQSIKOSKMAIDAQ
KLVVIVIVVPLLFKPIIGPKTKPVLDPKRNDFQSPFLVEKTLTHNTSMYKFLGLPH
ADVLGLPIGQHIVIKANINGKQITRSYTFSLDGTGKGFELVKSYPGTGNSKMGIC
ELKIGDSIQINGPRGNTHYERNCRSHGMIAGGTGIAPIQTMKAIAMDPRDTTKVSL
VFGNVHEEDILLKKEALVAMKPSQPKIVYLDSPREDTQMTGCVGXIITKDVIKHLP
AATMDNVQILICGPPAMVASVRRSTVDLGFRRSKPLSKMEDQVEVF"
order(1423. .1426,1428. .1432,1447. .1450)
polya_signal
ORIGIN

Alignment Scores:
Pred. No.: 540 Length: 1586
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x SCCYTB5R (1-1586)

Qy 1 GlyTyr***ValGluGlu 6
Db 1062 GGATATTCTGTGAAGAA 1079

RESULT 62
AF041083/c
LOCUS AF041083 1607 bp mRNA linear ROD 19-FEB-1998
DEFINITION Rattus norvegicus RoBo-1 mRNA, complete cds.
ACCESSION AF041083
VERSION AF041083.1 GI:2895562
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1607)
AUTHORS Noel,L.S., Champion,B.R., Holley,C.L., Simmons,C.J., Morris,D.C.,
Payne,J.A., Lean,J.M., Chambers,T.J., Zaman,G., Lanyon,L.E.,
Suva,L.J. and Miller,L.R.
TITLE RoBo-1, a novel member of the urokinase plasminogen activator
receptor/CD59/Ly-6/snake toxin family selectively expressed in rat
bone and growth plate cartilage
J. Biol. Chem. 273 (7), 3878-3883 (1998)
JOURNAL 98129789
MEDLINE 9461570
PUBMED
REFERENCE 2 (bases 1 to 1607)
AUTHORS Noel,L.S. and Miller,L.R.
TITLE Direct Submission
JOURNAL Submitted (05-JAN-1998) Molecular Pharmacology, Glaxo Wellcome
Research, 5 Moore Dr., Research Triangle Park, NC 27709, USA
FEATURES
source
1. .1607
/organism="Rattus norvegicus"


```
AUTHORS      Pogson,G.H. and Mesa,K.A.
TITLE        Positive darwinian selection at the pantophysin (pan I) locus in
JOURNAL      marine gadid fishes
PUBMED       Mol. Biol. Evol. 21 (1), 65-75 (2004)
AUTHORS      2 (bases 1 to 1806)
TITLE        Pogson,G.H. and Mesa,K.A.
JOURNAL      Direct Submission
SUBMITTED    (05-MAY-2003) Department of Ecology and Evolutionary
BIOLOGY,    University of California, Santa Cruz, Earth & Marine
SCIENCES     Bldg., Santa Cruz, CA 95064, USA
FEATURES     Location/Qualifiers
source       1..1806
            /organism="Trisopterus esmarkii"
            /mol_type="genomic DNA"
            /db_xref="taxon:80723"
            <1..>1775
            /gene="Pan I"
            join(<534..652,753..948,1211..1399,1653..>1775)
            /gene="Pan I"
            /product="pantophysin"
            join(<534..652,753..948,1211..1399,1653..>1775)
            /gene="Pan I"
            /codon_start=1
            /product="pantophysin"
            /protein_id="AA062774.1"
            /db_xref="GI:34148533"
            /translation="IFSIFAFATTGGYSGTTSINVKCPASDSQBSISANFNYPFLMHO
SYVPPDCNGTTFETFLIGNHSSSAEFFVSGVLSFLYSIASLVLYLGFQHYRKTSG
PVLDFVTGALAFMLVSSAWGKGLTDVXFPTNIVALIVACRSLGCTAGSVFS
MGRNASVMFGLNLILWGSNCWFYKTFPHKSNQPEAVEGRGPPT"

ORIGIN
Alignment Scores:
Pred. No.:      626      Length:      1806
Score:          27.00    Matches:      5
Percent Similarity: 83.33%  Conservative: 0
Best Local Similarity: 83.33%  Mismatches: 1
Query Match:     96.43%    Indels:      0
DB:              5        Gaps:        0

US-10-030-194A-5 (1-6) x AY292497 (1-1806)

Qy      1 GlyTyr***ValGluGlu 6
Db      885 GGCTATGCTGTAGAGGAA 868

RESULT 66
AY292496/c      1835 bp      DNA      linear      VRT 11-MAR-2004
LOCUS          Micromesistius poutassou pantophysin (Pan I) gene, partial cds.
DEFINITION     Micromesistius poutassou
ACCESSION      AY292496
VERSION        AY292496.1 GI:34148530
KEYWORDS       .
SOURCE         Micromesistius poutassou (blue whiting)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Actinopterygii; Neopterygii; Teleostei; Euteleostei;
               Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae;
               Micromesistius.
REFERENCE      1 (bases 1 to 1835)
AUTHORS        Pogson,G.H. and Mesa,K.A.
TITLE          Positive darwinian selection at the pantophysin (pan I) locus in
               marine gadid fishes
JOURNAL        Mol. Biol. Evol. 21 (1), 65-75 (2004)
PUBMED         12949133
AUTHORS        2 (bases 1 to 1835)
TITLE          Pogson,G.H. and Mesa,K.A.
JOURNAL        Direct Submission
SUBMITTED      (05-MAY-2003) Department of Ecology and Evolutionary
BIOLOGY,       University of California, Santa Cruz, Earth & Marine
SCIENCES       Bldg., Santa Cruz, CA 95064, USA
FEATURES       Location/Qualifiers
```

```
1..1835
/organism="Micromesistius poutassou"
/mol_type="genomic DNA"
/db_xref="taxon:81636"
<1..>1808
/gene="Pan I"
join(<499..617,727..922,1223..1411,1686..>1808)
/gene="Pan I"
/product="pantophysin"
join(<499..617,727..922,1223..1411,1686..>1808)
/gene="Pan I"
/codon_start=1
/product="pantophysin"
/protein_id="AA062773.1"
/db_xref="GI:34148531"
/translation="IFSIFAFATTGGYSGTTSINVCPGNDSOBITANFNYPFLMHO
SYVPPDCNGTTFETFLIGNHSSSAEFFVSGVLSFLYSIASLVLYLGFQHYRKTSG
PVLDFVTGALAFMLVSSAWGKGLTDVXFPTNIVALIVACRSLGCTAGSVPH
MGRNASVIFGLNLILWGSNCWFYKTFPHKSTNQPEDAEGRGPTT"

ORIGIN
Alignment Scores:
Pred. No.:      638      Length:      1835
Score:          27.00    Matches:      5
Percent Similarity: 83.33%  Conservative: 0
Best Local Similarity: 83.33%  Mismatches: 1
Query Match:     96.43%    Indels:      0
DB:              5        Gaps:        0

US-10-030-194A-5 (1-6) x AY292496 (1-1835)

Qy      1 GlyTyr***ValGluGlu 6
Db      859 GGCTATGCTGTAGAGGAA 842

RESULT 67
AY292490/c      1865 bp      DNA      linear      VRT 11-MAR-2004
LOCUS          Boreogadus saida pantophysin (Pan I) gene, partial cds.
DEFINITION     Boreogadus saida
ACCESSION      AY292490
VERSION        AY292490.1 GI:34148518
KEYWORDS       .
SOURCE         Boreogadus saida
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Actinopterygii; Neopterygii; Teleostei; Euteleostei;
               Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae;
               Boreogadus.
REFERENCE      1 (bases 1 to 1865)
AUTHORS        Pogson,G.H. and Mesa,K.A.
TITLE          Positive darwinian selection at the pantophysin (pan I) locus in
               marine gadid fishes
JOURNAL        Mol. Biol. Evol. 21 (1), 65-75 (2004)
PUBMED         12949133
AUTHORS        2 (bases 1 to 1865)
TITLE          Pogson,G.H. and Mesa,K.A.
JOURNAL        Direct Submission
SUBMITTED      (05-MAY-2003) Department of Ecology and Evolutionary
BIOLOGY,       University of California, Santa Cruz, Earth & Marine
SCIENCES       Bldg., Santa Cruz, CA 95064, USA
FEATURES       Location/Qualifiers
source       1..1865
            /organism="Boreogadus saida"
            /mol_type="genomic DNA"
            /db_xref="taxon:44932"
            <1..>1859
            /gene="Pan I"
            join(<480..598,722..923,1300..1488,1737..>1859)
            /gene="Pan I"
            /product="pantophysin"
            join(<480..598,722..923,1300..1488,1737..>1859)
            /gene="Pan I"
            /codon_start=1
            /product="pantophysin"
```

```
/protein_id="AAQ62767.1"
/db_xref="GI:34148519"
/translation="IFSIFAFATGGGSGTTSSINVCKSGVNOEIIYASFNFYFRLMQH
PKYVCKNGTTEESFLIGNYSSAEFFVSGISLVLSYLVLGFEHLVRKTS
RGPVVDLFTVLAALFLWLVSSAWKGLTDVKWATSPTSIVALSDVCRSDCTAGSV
PHMGLNASVIFGLNLLWGSNCWFYIYKTFPHKSNQPEDAEGRPPT"

ORIGIN
exon      151..277
/gene="gpaB"
/number=1
intron    278..356
/gene="gpaB"
/number=1
exon      357..702
/gene="gpaB"
/number=2
intron    703..758
/gene="gpaB"
/number=2
exon      759..887
/gene="gpaB"
/number=3
intron    888..950
/gene="gpaB"
/number=3
exon      951..1080
/gene="gpaB"
/number=4
intron    1081..1136
/gene="gpaB"
/number=4
exon      1137..1375
/gene="gpaB"
/number=5
intron    1376..1444
/gene="gpaB"
/number=5
exon      1445..1544
/gene="gpaB"
/number=6

ORIGIN
Alignment Scores:
Pred. No.: 661 Length: 1894
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x AFU277436 (1-1894)
Qy 1 GlyTyr***ValGluGlu 6
Db 416 GGCTACACAGTGGAGGAA 433

RESULT 69
AX079054/c
LOCUS AX079054 1896 bp DNA linear PAT 22-FEB-2001
DEFINITION Sequence 8 from Patent WO0107629.
ACCESSION AX079054
VERSION AX079054.1 GI:13158628
KEYWORDS Pseudomonas putida
SOURCE Pseudomonas putida
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1
AUTHORS Fraser,C.M., Venter,C., Tuemmler,B., Hoheisel,J., Duesterhoeft,A.,
Hilbert,H., Timmis,K.N., Moore,E., Straetz,M. and Heim,S.
TITLE Dna sequences which code oxygenases
JOURNAL Patent: WO 0107629-A 8 01-FEB-2001;
The Institute for Genomic Research (US) ; Quiagen GmbH (DE) ;
Gesellschaft fuer Biotechnologische Forschung mbH (GBF) (DE) ;
Deutsches Krebsforschungszentrum (DE) ; Medizinische Hochschule
Hannover (DE)
FEATURES
source
1..1896
/mol_type="unassigned DNA"
```

```
/protein_id="AAQ62767.1"
/db_xref="GI:34148519"
/translation="IFSIFAFATGGGSGTTSSINVCKSGVNOEIIYASFNFYFRLMQH
PKYVCKNGTTEESFLIGNYSSAEFFVSGISLVLSYLVLGFEHLVRKTS
RGPVVDLFTVLAALFLWLVSSAWKGLTDVKWATSPTSIVALSDVCRSDCTAGSV
PHMGLNASVIFGLNLLWGSNCWFYIYKTFPHKSNQPEDAEGRPPT"

ORIGIN
exon      151..277
/gene="gpaB"
/number=1
intron    278..356
/gene="gpaB"
/number=1
exon      357..702
/gene="gpaB"
/number=2
intron    703..758
/gene="gpaB"
/number=2
exon      759..887
/gene="gpaB"
/number=3
intron    888..950
/gene="gpaB"
/number=3
exon      951..1080
/gene="gpaB"
/number=4
intron    1081..1136
/gene="gpaB"
/number=4
exon      1137..1375
/gene="gpaB"
/number=5
intron    1376..1444
/gene="gpaB"
/number=5
exon      1445..1544
/gene="gpaB"
/number=6

ORIGIN
Alignment Scores:
Pred. No.: 661 Length: 1894
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x AY292490 (1-1865)
Qy 1 GlyTyr***ValGluGlu 6
Db 860 GGCTATGCTGTAGAGGAA 843

RESULT 68
AFU277436
LOCUS AFU277436 1894 bp DNA linear PLN 09-JUL-2003
DEFINITION Aspergillus fumigatus gpaB gene for guanine nucleotide binding
protein alpha subunit, exons 1-6.
ACCESSION AJ277436
VERSION AJ277436.1 GI:27524349
KEYWORDS gpaB gene; guanine nucleotide binding protein alpha subunit.
SOURCE Aspergillus fumigatus
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1
AUTHORS Liebmann,B., Gattung,S., Jahn,B. and Brakhage,A.A.
TITLE cAMP signaling in Aspergillus fumigatus involved in the
regulation of the virulence gene pksP and in defense against killing
by macrophages
JOURNAL Mol. Genet. Genomics 269 (3), 420-435 (2003)
MEDLINE 22719843
PUBMED 12734751
REFERENCE 2 (bases 1 to 1894)
AUTHORS Liebmann,B.
TITLE Direct Submision
JOURNAL Submitted (14-APR-2000) Liebmann B., Institute for Microbiology and
Genetics, Technical University, Schnittpahstr. 10, Darmstadt,
64287, GERMANY
FEATURES
source
1..1894
/organism="Aspergillus fumigatus"
/mol_type="genomic DNA"
/strain="ATCC46645"
/db_xref="taxon:5085"
151..1544
/gene="gpaB"
join(151..277,357..702,759..887,951..1080,1137..1375,
1445..1544)
/gene="gpaB"
/codon_start=1
/product="guanine nucleotide binding protein alpha
subunit"
/db_xref="GI:27524350"
/db_xref="GOA:Q8TFX6"
/db_xref="TrEMBL:Q8TFX6"
/translation="MGSCFSTESPDAEQKRSQAIDRRLEDSRRRLRECKILLGS
GESGKTVKQMKIYHQNGYTVBELALYRLTIYKNLLDCAKALIGAYHQFNQPTSRK
VREFIEFLSEYNIIDPNTPLDPRIGDAVTYLVNDPCTSTALEHQNFYLMDSAPYFF
BEAKRIAPDYPNVADVLARTKTGTGYETFTMGOLS IHMFVGGORSERKKWIHC
PENTSIIFCVALSEYDOVLEESNQRMESLVLFDSVYNSRPFMTSIIILFNKVD
LFRQKLPRSPUSNYFPDVGNDVNRRAKYLRLFRNQVNRALHLYLPHLTQATDTTNI
RLVFAAVKETILQNKDSGL"
```

/db_xref="taxon:303"

ORIGIN

Alignment Scores:
Pred. No.: 662 Length: 1896
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x AX079054 (1-1896)

Qy 1 GlyTyr***ValGluGlu 6

Db 51 GGCTACAGTGTGGAAGAG 34

RESULT 70

AY292485/c 1932 bp DNA linear VRT 11-MAR-2004
LOCUS Arctogadus glacialis pantophysin (Pan I) gene, partial cds.
DEFINITION Arctogadus glacialis pantophysin (Pan I) gene, partial cds.
ACCESSION AY292485
VERSION AY292485.1 GI:34148508
KEYWORDS
SOURCE Arctogadus glacialis (Arctic cod)
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Arctogadus.
TITLE 1 (bases 1 to 1932)
JOURNAL Pogson, G.H. and Mesa, K.A.
PUBMED Positive darwinian selection at the pantophysin (pan I) locus in
12949133 marine gadid fishes
Mol. Biol. Evol. 21 (1), 65-75 (2004)

REFERENCE 2 (bases 1 to 1932)
AUTHORS Pogson, G.H. and Mesa, K.A.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2003) Department of Ecology and Evolutionary
Biology, University of California, Santa Cruz, Earth & Marine
Sciences Bldg., Santa Cruz, CA 95064, USA

FEATURES
source
1..1932
/organism="Arctogadus glacialis"
/mol_type="genomic DNA"
/db_xref="taxon:185735"

gene
1..1884
/gene="Pan I"

mRNA
join(<510..628,751..946,1325..1513,1762..1884)
/gene="Pan I"

CDS
join(<510..628,751..946,1325..1513,1762..1884)
/gene="Pan I"

/product="pantophysin"
/codon_start=1
/product="pantophysin"

/protein_id="AA062762.1"
/db_xref="GI:34148509"

/translation="IFSIIFATTTGGFSGTTSINVQKGSVNQISAFNYPPFLMQH
PKYVPCDKNGTTEFLIGNYSSSAEFFVIGVLFSYIASLVLYLGFHLYRKTSG
PVVDFVTAALAFVLWSSAWGKGLTDVKWATPTSIVALSDACRSDGNCAGTGVPH
MGRNLASVIFGFLNLILWGSNCWFIYKTPPHKSANQPEDAEGRGPT"

ORIGIN

Alignment Scores:
Pred. No.: 676 Length: 1932
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x AY292485 (1-1932)

Qy 1 GlyTyr***ValGluGlu 6

Db 883 GGCTATGCTGTAGAGAA 866

RESULT 71

AX595606 2000 bp DNA linear PAT 14-FEB-2003
LOCUS Sequence 1260 from Patent EP1258494.
DEFINITION AX595606
ACCESSION AX595606
VERSION AX595606.1 GI:28396909
KEYWORDS
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM

REFERENCE
AUTHORS Saccharomyces cerevisiae
Marzioch, M., Schultz, J. and Superti-Furga, G.
TITLE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
JOURNAL Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
Patent: EP 1258494-A 1260 20-NOV-2002;
CELLZOME AG (DE)
FEATURES
source
1..2000
/organism="Saccharomyces cerevisiae"
/mol_type="unassigned DNA"
/db_xref="taxon:4932"

ORIGIN

Alignment Scores:
Pred. No.: 703 Length: 2000
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x AX595606 (1-2000)

Qy 1 GlyTyr***ValGluGlu 6

Db 1693 GGTATACGGTAGAAGAA 1710

RESULT 72

AX819744 2000 bp DNA linear PAT 10-DEC-2003
LOCUS Sequence 1494 from Patent EP1338608.
DEFINITION AX819744
ACCESSION AX819744
VERSION AX819744.1 GI:39724034
KEYWORDS
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM

REFERENCE
AUTHORS Saccharomyces cerevisiae
Bauer, A., Gavin, A.C., Superti-Furga, G., Kuester, B., Schultz, J.,
Marzioch, M., Grandi, P., Krause, R., Kruse, U., Merino, A., Bauch, A.,
Michon, A.M., Leutwein, C. and Rick, J.
TITLE Protein complexes and methods for their use
JOURNAL Patent: EP 1338608-A 1494 27-AUG-2003;
CELLZOME AG (DE)
FEATURES
source
1..2000
/organism="Saccharomyces cerevisiae"
/mol_type="unassigned DNA"
/db_xref="taxon:4932"

ORIGIN

Alignment Scores:
Pred. No.: 703 Length: 2000
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x AX595606 (1-2000)

Qy 1 GlyTyr***ValGluGlu 6

Db 1693 GGTATACGGTAGAAGAA 1710

RESULT 72

AX819744 2000 bp DNA linear PAT 10-DEC-2003
LOCUS Sequence 1494 from Patent EP1338608.
DEFINITION AX819744
ACCESSION AX819744
VERSION AX819744.1 GI:39724034
KEYWORDS
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM

REFERENCE
AUTHORS Saccharomyces cerevisiae
Bauer, A., Gavin, A.C., Superti-Furga, G., Kuester, B., Schultz, J.,
Marzioch, M., Grandi, P., Krause, R., Kruse, U., Merino, A., Bauch, A.,
Michon, A.M., Leutwein, C. and Rick, J.
TITLE Protein complexes and methods for their use
JOURNAL Patent: EP 1338608-A 1494 27-AUG-2003;
CELLZOME AG (DE)
FEATURES
source
1..2000
/organism="Saccharomyces cerevisiae"
/mol_type="unassigned DNA"
/db_xref="taxon:4932"

ORIGIN

Alignment Scores:
Pred. No.: 703 Length: 2000
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0

```
DB: 6 Gaps: 0
US-10-030-194A-5 (1-6) x AX819744 (1-2000)
QY 1 GlyTyr***ValGluGlu 6
DB 1693 GGTATACGGTAGAGAA 1710

RESULT 73
AX830774 2000 bp DNA linear PAT 12-DEC-2003
LOCUS
DEFINITION Sequence 1494 from Patent WO03072602.
ACCESSION AX830774
VERSION AX830774.1 GI:39839732
KEYWORDS
SOURCE
ORGANISM Saccharomyces cerevisiae (baker's yeast)
REFERENCE 1 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
AUTHORS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Bauer, A., Gavin, A.C., Superti-Furga, G., Kuester, B., Schultz, J.,
Marzioch, M., Grandi, P., Krause, R., Kruse, U., Merino, A., Bauch, A.,
Michon, A.M., Leutwein, C. and Rick, J.
TITLE Protein complexes and methods for their use
JOURNAL Patent: WO 03072602-A 1494 04-SEP-2003;
CELLZONE AG (DE)
FEATURES
source
1..2000
/organism="Saccharomyces cerevisiae"
/mol_type="unassigned DNA"
/db_xref="taxon:4932"

ORIGIN
Alignment Scores: 703 Length: 2000
Pred. No.: 27.00 Matches: 5
Score: 83.33% Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.43% Gaps: 0
DB: 6

US-10-030-194A-5 (1-6) x AX830774 (1-2000)
QY 1 GlyTyr***ValGluGlu 6
DB 1693 GGTATACGGTAGAGAA 1710

RESULT 74
AX927178 2036 bp DNA linear PAT 19-DEC-2003
LOCUS
DEFINITION Sequence 99 from Patent WO03085115.
ACCESSION AX927178
VERSION AX927178.1 GI:40248025
KEYWORDS
SOURCE
ORGANISM Nicotiana tabacum (common tobacco)
REFERENCE 1 Nicotiana tabacum
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
JOURNAL asterids; lamids; Solanales; Solanaceae; Nicotiana.
FEATURES
source
1..2036
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

ORIGIN
Alignment Scores: 718 Length: 2036
Pred. No.: 27.00 Matches: 5
Score: 83.33% Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.43% Gaps: 0
DB: 6

US-10-030-194A-5 (1-6) x AX834957 (1-2104)
QY 1 GlyTyr***ValGluGlu 6
DB 2058 GGATATGCAGTAGAGAA 2041

Search completed: November 3, 2004, 16:30:34
Job time : 2166 secs
```

```
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x AX927178 (1-2036)
QY 1 GlyTyr***ValGluGlu 6
DB 1929 GGTATACAGTAGAGAA 1912

RESULT 75
AX834957 2104 bp DNA linear PAT 15-DEC-2003
LOCUS
DEFINITION Sequence 2081 from Patent EP1347046.
ACCESSION AX834957
VERSION AX834957.1 GI:39921092
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J.I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and
Masuho, Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 2081 24-SEP-2003;
FEATURES Research Association for Biotechnology (JP)
source
1..2104
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 745 Length: 2104
Pred. No.: 27.00 Matches: 5
Score: 83.33% Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.43% Gaps: 0
DB: 6
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2004, 11:22:27 ; Search time 314 Seconds

(without alignments)
100.307 Million cell updates/sec

Title: US-10-030-194A-5

Perfect score: 28

Sequence: 1 GYXVEE 6

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODE=frame+p2n.model -DEV=xlp

-Q=/cgn2_1/USPRO.spool_p/US10030194/runat_01112004_184846_28945/app_query.fasta_1.398

-DB=N Geneseq_23Sep04 -QFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0

-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi

-LIST=1000 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75

-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10030194 @CGN 1 1 1063 @runat_01112004_184846_28945 -NCPU=6 -ICPU=3

-NO_WMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_23Sep04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	96.4	318	6	ABN76795 Human red
2	27	96.4	330	2	AaQ04657 Sequence
3	27	96.4	330	2	AaQ04594 Sequence
4	27	96.4	330	2	AaQ068744 CHA255 li
5	27	96.4	384	8	ABX64662 Human gen
6	27	96.4	400	5	ABV07265 Human pro

7	27	96.4	417	9	ACH17756
8	27	96.4	431	5	ABV37205
9	27	96.4	460	5	ABA13307
10	27	96.4	479	6	ABK77009
11	27	96.4	490	6	AB199741
12	27	96.4	498	6	ABK62521
13	27	96.4	498	10	ADB56048
14	27	96.4	533	10	ADC75197
15	27	96.4	535	6	ABK73924
16	27	96.4	565	5	ABV59077
17	27	96.4	696	10	ACF68403
18	27	96.4	708	2	AAZ15431
19	27	96.4	747	4	AAK90974
20	27	96.4	747	5	AAK32009
21	27	96.4	747	6	ABN90364
22	27	96.4	747	11	ADJ15277
23	27	96.4	834	8	ACA18963
24	27	96.4	849	10	ADC91913
25	27	96.4	963	4	AAH31919
26	27	96.4	987	5	AAK42359
27	27	96.4	987	6	ABK16635
28	27	96.4	987	6	ABZ43050
29	27	96.4	987	6	ABK68526
30	27	96.4	987	6	ABK37645
31	27	96.4	987	12	ADG83335
32	27	96.4	1001	12	ADQ35088
33	27	96.4	1002	6	ABK97222
34	27	96.4	1002	10	AAK60343
35	27	96.4	1035	4	AAK51422
36	27	96.4	1035	8	ACA19457
37	27	96.4	1050	11	ABD12369
38	27	96.4	1065	10	ADC91898
39	27	96.4	1239	6	ABN67444
40	27	96.4	1252	12	ADO40238
41	27	96.4	1269	12	ADJ39779
42	27	96.4	1296	12	ADQ80770
43	27	96.4	1347	5	AAK81888
44	27	96.4	1347	10	ADD46198
45	27	96.4	1347	10	ADG62406
46	27	96.4	1387	10	ADC85802
47	27	96.4	1497	4	AAF61015
48	27	96.4	1635	4	AAK51954
49	27	96.4	1659	8	ACF74491
50	27	96.4	1662	4	AAK54463
51	27	96.4	1662	8	ACA20007
52	27	96.4	1857	5	AAK81886
53	27	96.4	1896	4	AAF26350
54	27	96.4	2000	10	ACC61239
55	27	96.4	2000	10	ADK63243
56	27	96.4	2036	10	ADF38040
57	27	96.4	2094	4	AAK07614
58	27	96.4	2104	11	ADM03396
59	27	96.4	2289	2	AAK27286
60	27	96.4	2289	6	AAK35185
61	27	96.4	2289	8	ABX14889
62	27	96.4	2547	11	ABD12667
63	27	96.4	2682	6	AAK62441
64	27	96.4	2964	11	ABD12308
65	27	96.4	3645	5	ABV24673
66	27	96.4	3645	5	ABV28517
67	27	96.4	3645	5	ABV22692
68	27	96.4	3645	5	ABV29547
69	27	96.4	3962	6	ABL58963
70	27	96.4	4217	8	ACA21975
71	27	96.4	5728	5	AAK81892
72	27	96.4	5925	8	ABL33576
73	27	96.4	6570	8	ACF34525
74	27	96.4	6571	6	ABL69964
75	27	96.4	6571	6	ABK83708
76	27	96.4	6577	10	ADB47395
77	27	96.4	8790	10	ADD13780
78	27	96.4	13321	2	AAV74315
79	27	96.4	17000	6	AAK40299

ACH17756	Human adu
ABV37205	Human pro
ABA13307	Human ner
ABK77009	Bacillus
AB199741	Mouse isc
ABK62521	Rat seque
ADB56048	Toxicity-
ADC75197	T harzian
ABK73924	Bacillus
ABV59077	Human pro
ACF68403	Photorhab
AZ15431	Human gen
AAK90974	Human dig
AAK32009	Human liv
ABN90364	Human liv
ADJ15277	Human liv
ACA18963	Prokaryot
ADC91913	E. faeciu
AAH31919	Human Olf
AAK42359	Human CDN
ABK16635	Human G-c
ABZ43050	Human GPC
ABK68526	Human DNA
ABK37645	DNA encod
ADG83335	Human Olf
ADQ35088	Human KCh
ABK97222	Human G-p
AAK60343	Human G-p
AAK51422	Pseudomon
ACA19457	Prokaryot
ABD12369	Pseudomon
ADC91898	E. faeciu
ABN67444	Streptoco
ADO40238	Streptoco
ADJ39779	Plant CDN
ADQ80770	Porcine e
AAK81888	DNA encod
ADD46198	Rat gene
ADG62406	Rat gene
ADC85802	Human GPC
AAF61015	P. putida
AAK51954	Staphyloc
ACF74491	Staphyloc
AAK54463	Staphyloc
ACA20007	Prokaryot
AAK81886	DNA encod
AAF26350	P. putida
ACC61239	Gene sequ
ADK63243	Disease t
ADF38040	Synchroni
AAK07614	Human sec
ADM03396	Human CDN
AAK27286	Desulfuro
AAK35185	Desulfuro
ABX14889	DNA encod
ABD12667	Pseudomon
AAK62441	cDNA sequ
ABD12308	Pseudomon
ABV24673	Human pro
ABV28517	Human pro
ABV22692	Human pro
ABV29547	Human pro
ABL58963	Human tum
ACA21975	Prokaryot
AAK81892	DNA encod
ABL33576	Human inn
ACF34525	Gene enco
ABL69964	Pancreeas
ABK83708	Human CDN
ADB47395	Human CDN
ADD13780	Plasmid p
AAV74315	Staphyloc
AAK40299	Caspase 6

80	27	96.4	17276	8	ACA64723	Aca64723 Streptoco	153	26	92.9	324	10	ADC92548	Adc92548 E. faeciu
81	27	96.4	17276	10	ADFA3363	Adf43363 Streptoco	154	26	92.9	336	3	AACT75546	Aac75546 Human ORF
82	27	96.4	25020	12	ADOA0235	Ado40235 S. agalac	155	26	92.9	336	6	ABN76782	Abn76782 Human iso
c 83	27	96.4	33954	3	AAZ93712	Aaz93712 F-box pro	156	26	92.9	343	4	AAI27264	Aai27264 Probe #17
84	27	96.4	96598	9	ADOA02609	Ado2609 Mouse Nfk	157	26	92.9	343	4	ABA75551	Ab75551 Human foe
85	27	96.4	96598	10	ADB72347	Adb72347 Mouse Nfk	158	26	92.9	343	4	AAI56124	Aai56124 Probe #24
86	27	96.4	96598	10	AD82941	Ad82941 Mouse Nfk	159	26	92.9	343	4	ABA40164	Ab40164 Probe #18
87	27	96.4	96598	10	AD85857	Ad85857 Mouse Nfk	160	26	92.9	343	4	AAK50184	Aak50184 Human bon
88	27	96.4	102644	10	ACF65378	Acf65378 Photorhab	161	26	92.9	343	4	AAK24126	Aak24126 Human bra
c 89	27	96.4	110000	6	ABN71527_11	Continuation (12 o	162	26	92.9	343	4	ABSA49815	Ab849815 Human liv
90	27	96.4	110000	6	ABA90521_09	Continuation (10 o	c 163	26	92.9	344	5	ABV35201	Abv35201 Human pro
91	27	96.4	110000	10	ACF67367_11	Continuation (12 o	163	26	92.9	353	4	AAK60687	Aak60687 Human imm
92	27	96.4	110000	11	ADM27081_00	Adm27081 Hyperther	c 164	26	92.9	353	4	AAAF11785	Aaf11785 Aspergill
c 93	27	96.4	110000	12	ADNA6845_05	Continuation (6 of	c 165	26	92.9	358	3	AAH69151	Aah69151 Human cer
c 94	27	96.4	110000	12	ADNA6845_06	Continuation (7 of	c 166	26	92.9	360	4	AAH69151	Aah69151 Human cer
95	27	96.4	110000	12	ADNA7591_14	Continuation (15 o	c 167	26	92.9	372	3	AAAC65302	Aac65302 Arabidops
c 96	27	96.4	110000	12	ADNA6123_05	Continuation (6 of	c 168	26	92.9	382	6	ABQ85282	Abq85282 Arabidops
c 97	27	96.4	110000	12	ADNA46123_06	Continuation (7 of	c 169	26	92.9	383	3	AAAC31586	Aac31586 Human sec
98	27	96.4	110000	12	ADNA7209_14	Continuation (15 o	c 170	26	92.9	387	10	ADH82280	Adh82280 Enterococ
c 99	27	96.4	110000	12	ADNA4646_05	Continuation (6 of	c 171	26	92.9	394	4	AAI20870	Aai20870 Probe #10
c 100	27	96.4	110000	12	ADNA4646_06	Continuation (7 of	c 172	26	92.9	394	4	AAI46108	Aai46108 Probe #14
101	27	96.4	110000	12	ADNA7960_14	Continuation (13 o	c 173	26	92.9	394	4	ABR48057	Ab48057 Human bre
c 102	27	96.4	110000	12	ADFA3927_2	Continuation (5 of	c 174	26	92.9	394	4	ABA33032	Ab33032 Probe #11
c 103	27	96.4	349980	5	AHA41225	Aha41225 Pyrococcu	c 175	26	92.9	394	4	AAK40090	Aak40090 Human bon
c 104	26	92.9	65	6	ABN55356	Abn55356 Mouse spl	c 176	26	92.9	394	4	AAK14361	Aak14361 Human bra
105	26	92.9	107	3	AAAC32334	Aac32334 Human sec	c 177	26	92.9	394	4	ABR339674	Ab339674 Human liv
106	26	92.9	150	4	AAI19718	Aai19718 Probe #96	c 178	26	92.9	402	5	ABV35075	Abv35075 Human pro
107	26	92.9	150	4	ABA64744	Ab64744 Human foe	c 179	26	92.9	403	3	AAAC11741	Aac11741 Human sec
108	26	92.9	150	4	AAI44913	Aai44913 Probe #13	180	26	92.9	404	3	AAAC24351	Aac24351 Human sec
109	26	92.9	150	4	ABA46863	Ab46863 Human bre	181	26	92.9	406	12	ADL85807	Adl85807 DNA up-re
110	26	92.9	150	4	ABA31866	Ab31866 Probe #10	182	26	92.9	406	12	ADL85807	Adl85807 DNA up-re
111	26	92.9	150	4	AAK38909	Aak38909 Human bon	183	26	92.9	407	12	ACH92706	Ach92706 Human gen
112	26	92.9	150	4	AAK13178	Aak13178 Human bra	184	26	92.9	408	4	AAI26760	Aai26760 Probe #16
113	26	92.9	150	4	ABR33967	Ab33967 Human liv	185	26	92.9	408	4	ABA75009	Ab75009 Human foe
114	26	92.9	150	5	AAI05437	Aai05437 Probe #54	186	26	92.9	408	4	AAI55537	Aai55537 Probe #24
115	26	92.9	150	6	ABR12991	Ab12991 Human gen	187	26	92.9	408	4	ABA39697	Ab39697 Probe #18
116	26	92.9	150	2	AAQ050198	Aaq050198 RGAE part	188	26	92.9	408	4	AAK23520	Aak23520 Human bra
117	26	92.9	165	4	AAI25367	Aai25367 Probe #15	189	26	92.9	408	4	ABR49277	Ab49277 Human liv
118	26	92.9	165	4	ABA71364	Ab71364 Human foe	190	26	92.9	410	5	ABA20968	Ab20968 Human ner
119	26	92.9	165	4	AAI51622	Aai51622 Probe #20	191	26	92.9	410	5	ABA20967	Ab20967 Human ner
120	26	92.9	165	4	AAK437611	Aak437611 Probe #16	192	26	92.9	430	12	ACH85177	Ach85177 Human gen
121	26	92.9	165	4	AAK45696	Aak45696 Human bon	193	26	92.9	431	12	ADP61834	Adp61834 Soybean c
122	26	92.9	165	4	AAK19677	Aak19677 Human bra	c 194	26	92.9	435	4	AAAL34969	Aal34969 Human mus
123	26	92.9	165	4	ABR45385	Ab45385 Human liv	c 195	26	92.9	435	8	ABX57957	Abx57957 cDNA enco
124	26	92.9	165	6	ABR19971	Ab19971 Human gen	c 196	26	92.9	435	12	ADH10451	Adh10451 Codon opt
125	26	92.9	165	12	ACH80878	Ach80878 Human gen	c 197	26	92.9	435	12	ADJ27684	Adj27684 Human mus
126	26	92.9	170	3	AAAC29351	Aac29351 Human sec	198	26	92.9	437	12	ACH85078	Ach85078 Human gen
127	26	92.9	170	12	ACH86164	Ach86164 Human gen	199	26	92.9	444	4	AAI26520	Aai26520 Probe #16
128	26	92.9	179	12	ACH86164	Ach86164 Human ner	200	26	92.9	444	4	ABA74750	Ab74750 Human foe
129	26	92.9	234	6	ABR69551	Ab69551 Novel mur	201	26	92.9	444	4	AAI55273	Aai55273 Probe #23
130	26	92.9	237	3	AAAC73048	Aac73048 Single nu	202	26	92.9	444	4	ABA39460	Ab39460 Probe #17
131	26	92.9	237	3	AAAC62502	Aac62502 Single nu	203	26	92.9	444	4	ABA39460	Ab39460 Probe #17
132	26	92.9	262	3	AAAC13358	Aac13358 Human sec	204	26	92.9	444	4	AAK23240	Aak23240 Human bra
133	26	92.9	263	3	AAAC03345	Aac03345 Human sec	205	26	92.9	444	4	ABSA49025	Ab49025 Human liv
c 134	26	92.9	270	2	AAQ84370	Aaq84370 Gemini vi	206	26	92.9	444	6	ABR22896	Ab22896 Human gen
c 135	26	92.9	270	6	ABR92185	Ab92185 Staphyloc	207	26	92.9	445	3	AAAX89447	Aax89447 Corn bran
136	26	92.9	273	4	AAI26826	Aai26826 Probe #16	208	26	92.9	452	5	ABA19784	Ab19784 Human ner
137	26	92.9	273	4	ABA75076	Ab75076 Human foe	209	26	92.9	456	4	AAI16167	Aai16167 Probe #61
138	26	92.9	273	4	AAI55610	Aai55610 Probe #24	210	26	92.9	456	4	ABAS8793	Abas8793 Human foe
139	26	92.9	273	4	ABA39760	Ab39760 Probe #18	211	26	92.9	456	4	AAI38489	Aai38489 Probe #71
140	26	92.9	273	4	AAK49719	Aak49719 Human bon	212	26	92.9	456	4	ABA27726	Ab27726 Probe #61
141	26	92.9	273	4	AAK23591	Aak23591 Human bra	213	26	92.9	456	4	AAK32674	Aak32674 Human bon
142	26	92.9	273	4	ABR49345	Ab49345 Human liv	214	26	92.9	456	4	AAK06941	Aak06941 Human bra
143	26	92.9	274	12	ACH88963	Ach88963 Human gen	215	26	92.9	456	4	ABR332391	Ab332391 Human liv
144	26	92.9	275	2	AAQ50196	Aaq50196 RGAI/3'-e	216	26	92.9	456	6	ABSO7468	Abso7468 Human gen
145	26	92.9	283	12	ACH86451	Ach86451 Human gen	217	26	92.9	461	9	ACH13578	Ach13578 Human adu
146	26	92.9	307	12	ACH80679	Ach80679 Human gen	218	26	92.9	463	5	ABV13967	Abv13967 Human pro
c 147	26	92.9	308	10	ADP85633	Adp85633 Human ade	219	26	92.9	464	3	AAAF16224	Aaf16224 Human pro
148	26	92.9	305	2	AAV52943	Aav52943 Human UNC	220	26	92.9	466	4	AAI10463	Aai10463 Probe #39
c 149	26	92.9	305	3	AAAC17969	Aac17969 Human sec	221	26	92.9	466	4	ABR52109	Ab52109 Human foe
150	26	92.9	309	12	ADK16100	Adk16100 Nanoarcha	222	26	92.9	466	4	AAI31716	Aai31716 Probe #40
151	26	92.9	317	2	AAQ61187	Aaq61187 Human bra	223	26	92.9	466	4	ABA21919	Ab21919 Probe #38
152	26	92.9	321	3	AAAC65320	Aac65320 Arabidops	224	26	92.9	466	4	AAK25834	Aak25834 Human bon
							225	26	92.9	466	4	AAK00391	Aak00391 Human bra

226	26	92.9	466	4	ABS25422	Abs25422 Human liv	299	26	92.9	570	4	AAI18122	Aai18122 Probe #80
227	26	92.9	466	5	AAI00399	Aai00399 Probe #39	300	26	92.9	570	4	AAI18129	Aai18129 Probe #80
228	26	92.9	466	6	ABS00415	Abs00415 Human gen	301	26	92.9	570	4	ABA63094	Aba63094 Human foe
C 229	26	92.9	467	9	ACH26223	Ach26223 Human adu	302	26	92.9	570	4	ABA63100	Aba63100 Human foe
C 230	26	92.9	478	3	ACA41205	Aca41205 Arabidops	303	26	92.9	570	4	AAI43127	Aai43127 Probe #11
C 231	26	92.9	479	2	AAQ84372	Aaq84372 Gemini vi	304	26	92.9	570	4	AAI43135	Aai43135 Probe #11
C 232	26	92.9	480	6	AB214808	Ab214808 Arabidops	305	26	92.9	570	4	ABA30345	Aba30345 Probe #88
C 233	26	92.9	486	6	ABQ69875	Abq69875 Listeria	306	26	92.9	570	4	AAK37302	Aak37302 Human bon
C 234	26	92.9	486	6	ABQ67957	Abq67957 Listeria	307	26	92.9	570	4	AAK37302	Aak37302 Human bon
C 235	26	92.9	495	6	ABK30383	Abk30383 Human G-p	308	26	92.9	570	4	AAK37295	Aak37295 Human bon
C 236	26	92.9	495	5	ADI73425	Adi73425 Human ova	309	26	92.9	570	4	AAK11526	Aak11526 Human bra
C 237	26	92.9	499	5	ADL38555	Adl38555 Human ova	310	26	92.9	570	4	AAK11517	Aak11517 Human bra
C 238	26	92.9	500	12	ADP93007	Adp93007 Cotton ex	311	26	92.9	570	4	ABS36972	Abs36972 Human liv
C 239	26	92.9	505	4	AAH72019	Aah72019 Human cer	312	26	92.9	570	4	ABS36965	Abs36965 Human liv
C 240	26	92.9	506	8	ABV76265	Abv76265 Tomato ye	313	26	92.9	570	6	ABS11285	Abs11285 Human gen
C 241	26	92.9	506	8	ABV76263	Abv76263 Tomato ye	314	26	92.9	570	6	ACH67173	Ach67173 Human gen
C 242	26	92.9	508	8	ABV76264	Abv76264 Tomato ye	315	26	92.9	571	12	ACH67173	Ach67173 Human gen
C 243	26	92.9	515	12	ACH72720	Ach72720 Human gen	316	26	92.9	573	6	ABK78055	Abk78055 Bacillus
C 244	26	92.9	519	4	AAI17621	Aai17621 Probe #75	317	26	92.9	575	2	AAV84491	Avv84491 Human sec
C 245	26	92.9	519	4	ABA62555	Aba62555 Human foe	318	26	92.9	575	2	ABA83274	Aba83274 Human sec
C 246	26	92.9	519	4	AAI42541	Aai42541 Probe #11	319	26	92.9	575	8	ADA33987	Ada33987 Human sec
C 247	26	92.9	519	4	ABA29883	Aba29883 Probe #83	320	26	92.9	575	9	ACH04775	Ach04775 Novel hum
C 248	26	92.9	519	4	AK367659	Aak367659 Human bon	321	26	92.9	575	9	ACD44585	Acd44585 Human cdn
C 249	26	92.9	519	4	AAK10911	Aak10911 Human bra	322	26	92.9	575	10	ADC73612	Adc73612 Human sec
C 250	26	92.9	519	4	ABS36429	Abs36429 Human liv	323	26	92.9	575	10	ADA56176	Ada56176 Gene enco
C 251	26	92.9	519	12	ACH90988	Ach90988 Human gen	324	26	92.9	576	4	AAI18635	Aai18635 Probe #85
C 252	26	92.9	522	5	ABV55091	Abv55091 Human pro	325	26	92.9	576	4	ABA63632	Aba63632 Human foe
C 253	26	92.9	523	4	AAI18097	Aai18097 Probe #80	326	26	92.9	576	4	AAI43747	Aai43747 Probe #12
C 254	26	92.9	523	4	ABA63064	Aba63064 Human foe	327	26	92.9	576	4	ABA30828	Aba30828 Probe #92
C 255	26	92.9	523	4	AAI43102	Aai43102 Probe #11	328	26	92.9	576	4	AK37873	Aak37873 Human bon
C 256	26	92.9	523	4	ABA30321	Aba30321 Probe #87	329	26	92.9	576	4	AAK12155	Aak12155 Human bra
C 257	26	92.9	523	4	AAK37268	Aak37268 Human bon	330	26	92.9	576	4	ABS37499	Abs37499 Human liv
C 258	26	92.9	523	4	AAK11489	Aak11489 Human bra	331	26	92.9	576	6	ABS11870	Abs11870 Human gen
C 259	26	92.9	523	4	ABS36938	Abs36938 Human liv	332	26	92.9	576	10	ABX56860	Abx56860 Arabidops
C 260	26	92.9	524	4	AK68272	Aak68272 Human imm	333	26	92.9	580	4	AAI19179	Aai19179 Probe #91
C 261	26	92.9	537	12	ACH67966	Ach67966 Human gen	334	26	92.9	580	4	ABA64186	Aba64186 Human foe
C 262	26	92.9	539	12	ACH75263	Ach75263 Human gen	335	26	92.9	580	4	AAI44334	Aai44334 Probe #13
C 263	26	92.9	542	4	AAI18255	Aai18255 Probe #81	336	26	92.9	580	4	ABA31336	Aba31336 Probe #98
C 264	26	92.9	542	4	ABA63234	Aba63234 Human foe	337	26	92.9	580	4	AAK38380	Aak38380 Human bon
C 265	26	92.9	542	4	AAI43304	Aai43304 Probe #11	338	26	92.9	580	4	AAK12665	Aak12665 Human bra
C 266	26	92.9	542	4	ABA30472	Aba30472 Probe #89	339	26	92.9	580	4	ABS37988	Abs37988 Human liv
C 267	26	92.9	542	4	AK37441	Aak37441 Human bon	340	26	92.9	583	6	ABS12434	Abs12434 Human gen
C 268	26	92.9	542	4	AAK11700	Aak11700 Human bra	341	26	92.9	584	4	AAI18234	Aai18234 Probe #81
C 269	26	92.9	542	4	ABS37125	Abs37125 Human liv	342	26	92.9	584	4	ABA63210	Aba63210 Human foe
C 270	26	92.9	543	4	AAI17299	Aai17299 Probe #72	343	26	92.9	584	4	AAI43276	Aai43276 Probe #11
C 271	26	92.9	543	4	ABA62212	Aba62212 Human foe	344	26	92.9	584	4	ABA30455	Aba30455 Probe #89
C 272	26	92.9	543	4	AAI42188	Aai42188 Probe #10	345	26	92.9	584	4	AK37416	Aak37416 Human bon
C 273	26	92.9	543	4	ABA29568	Aba29568 Probe #80	346	26	92.9	584	4	AKK11672	Aak11672 Human bra
C 274	26	92.9	543	4	AK36444	Aak36444 Human bon	347	26	92.9	584	4	ABS37099	Abs37099 Human liv
C 275	26	92.9	543	4	AAK10543	Aak10543 Human bra	348	26	92.9	584	6	ABS11410	Abs11410 Human gen
C 276	26	92.9	543	4	ABS36093	Abs36093 Human liv	349	26	92.9	590	12	ACH87680	Ach87680 Human gen
C 277	26	92.9	543	5	AS91277	As91277 DNA encod	350	26	92.9	590	12	ACH71421	Ach71421 Human gen
C 278	26	92.9	543	6	ABS10444	Abs10444 Human gen	351	26	92.9	600	12	ACH72408	Ach72408 Human gen
C 279	26	92.9	543	12	ADK16446	Adk16446 Nanoarcha	352	26	92.9	602	5	ABV60423	Abv60423 Human pro
C 280	26	92.9	547	12	ACH77287	Ach77287 Human gen	353	26	92.9	604	4	AAI19586	Aai19586 Probe #95
C 281	26	92.9	549	4	AAI17553	Aai17553 Probe #74	354	26	92.9	604	4	ABA64610	Aba64610 Human foe
C 282	26	92.9	549	4	ABA62486	Aba62486 Human foe	355	26	92.9	604	4	AAI44780	Aai44780 Probe #13
C 283	26	92.9	549	4	AAI42466	Aai42466 Probe #11	356	26	92.9	604	4	ABA46732	Aba46732 Human bre
C 284	26	92.9	549	4	ABA29818	Aba29818 Probe #82	357	26	92.9	604	4	ABA31737	Aba31737 Probe #10
C 285	26	92.9	549	4	AAK10838	Aak10838 Human bra	358	26	92.9	604	4	AK38780	Aak38780 Human bon
C 286	26	92.9	549	4	ABS36359	Abs36359 Human liv	359	26	92.9	604	4	AAK13049	Aak13049 Human bra
C 287	26	92.9	550	2	AAQ84376	Aaq84376 Tomato le	360	26	92.9	604	4	ABS38361	Abs38361 Human liv
C 288	26	92.9	550	2	AAQ84377	Aaq84377 Tomato le	361	26	92.9	604	5	AAI05306	Aai05306 Probe #52
C 289	26	92.9	550	10	ABX57188	Abx57188 Arabidops	362	26	92.9	604	6	ABS12856	Abs12856 Human gen
C 290	26	92.9	554	10	ABX61276	Abx61276 Arabidops	363	26	92.9	608	5	ABV44041	Abv44041 Human pro
C 291	26	92.9	556	10	ACD96589	Acd96589 Human col	364	26	92.9	608	5	ABV43921	Abv43921 Human pro
C 292	26	92.9	558	12	ACH79004	Ach79004 Human gen	365	26	92.9	621	6	ABQ43976	Abq43976 Oligonuc1
C 293	26	92.9	559	9	ACH40823	Ach40823 Human foe	366	26	92.9	621	6	ABQ43977	Abq43977 Oligonuc1
C 294	26	92.9	563	12	ACH71332	Ach71332 Human gen	367	26	92.9	622	6	ABQ15917	Abq15917 Oligonuc1
C 295	26	92.9	566	12	ACH66972	Ach66972 Human gen	368	26	92.9	622	6	ABQ15916	Abq15916 Oligonuc1
C 296	26	92.9	567	5	AA569569	Aas69569 DNA encod	369	26	92.9	624	3	AAK52149	Aak52149 Arabidops
C 297	26	92.9	567	5	AA581910	Aas81910 DNA encod	370	26	92.9	624	3	AAH70956	Aah70956 Human cer
C 298	26	92.9	567	5	ADL63706	Adl63706 Human ova	371	26	92.9	625	10	ADF69451	Adf69451 Tapezia y

372	26	92.9	642	10	ADB52003	Adb52003 Primary r	445	92.9	1154	3	AAC47523	Aac47523 Arabidops
373	26	92.9	646	6	ABQ20334	Abq20334 Oligonucl	446	92.9	1155	6	ABZ13425	Abz13425 Arabidops
374	26	92.9	646	6	ABQ20335	Abq20335 Oligonucl	447	92.9	1190	5	AAS66819	Aas66819 DNA encod
375	26	92.9	656	8	ABZ18396	Abz18396 Group III	448	92.9	1197	6	ABN69130	Abn69130 Streptoco
376	26	92.9	657	5	AAS66679	Aas66679 DNA encod	C 449	92.9	1200	11	ABD09864	Abd09864 Pseudomon
377	26	92.9	657	5	AAS66679	Aas66679 DNA encod	C 450	92.9	1214	3	ABD09864	Aac45611 Arabidops
378	26	92.9	662	10	ADK56024	Adk56024 Plant DNA	C 451	92.9	1215	5	AAS94499	Aas94499 DNA encod
379	26	92.9	663	6	ABQ21144	Abq21144 Oligonucl	C 452	92.9	1215	5	AAS79152	Aas79152 DNA encod
380	26	92.9	663	6	ABQ21145	Abq21145 Oligonucl	453	92.9	1216	3	AAC33721	Aac33721 Arabidops
381	26	92.9	668	5	ADL44939	Adl44939 Human ova	454	92.9	1220	3	AAC33193	Aac33193 Arabidops
382	26	92.9	675	2	AAX55798	Aax55798 Vegetativ	455	92.9	1224	4	AAF71883	AAF71883 Corynebac
383	26	92.9	691	8	ABV76266	Abv76266 Tomato ye	456	92.9	1239	10	ADP03438	Adp03438 Bacterial
384	26	92.9	711	4	ABL03427	AbL03427 Drosophil	457	92.9	1245	3	AAC42225	Aac42225 Arabidops
385	26	92.9	711	12	ADO07737	Ado07737 Fly polyn	458	92.9	1249	3	AAC45902	Aac45902 Arabidops
386	26	92.9	714	5	AAS66691	Aas66691 DNA encod	459	92.9	1252	2	AZ42053	Az42053 Human end
387	26	92.9	717	10	ADC32987	Adc32987 E. faeciu	460	92.9	1267	10	ADD13504	Add13504 C. glutam
388	26	92.9	721	3	ABA038701	AbA038701 Arabidops	461	92.9	1278	3	AAC32761	Aac32761 Arabidops
389	26	92.9	721	4	ABA08412	AbA08412 Human sec	462	92.9	1300	3	AAAF78551	AAAF78551 Plant SDF
390	26	92.9	753	5	AAS66815	Aas66815 DNA encod	463	92.9	1303	3	AAAF78453	AAAF78453 Plant SDF
391	26	92.9	771	2	AAT12178	Aat12178 Partial p	464	92.9	1330	5	AAAF72982	AAAF72982 DNA encod
392	26	92.9	782	4	AAC84576	Aac84576 Corn magn	465	92.9	1345	10	ADC86054	Adc86054 Human GPC
393	26	92.9	786	5	AAS66689	Aas66689 DNA encod	C 466	92.9	1383	5	AAS69699	Aas69699 DNA encod
394	26	92.9	845	5	ABZ20440	Abz20440 Oncofoeta	467	92.9	1383	6	ABN93022	Abn93022 Staphyloc
395	26	92.9	861	12	ACH91002	Ach91002 Human gen	468	92.9	1393	5	AAS66698	Aas66698 DNA encod
396	26	92.9	885	2	AAQ50199	Aaq50199 RGAE part	469	92.9	1395	3	AAA26979	Aaa26979 Human cor
397	26	92.9	890	6	ABQ14472	Abq14472 Oligonucl	470	92.9	1427	3	AAC42395	Aac42395 Arabidops
398	26	92.9	890	6	ABQ14473	Abq14473 Oligonucl	471	92.9	1434	5	AAS65653	Aas65653 DNA encod
399	26	92.9	894	5	AAS70828	Aas70828 DNA encod	C 472	92.9	1434	5	AAS86659	Aas86659 DNA encod
400	26	92.9	906	6	ABQ15942	Abq15942 Oligonucl	473	92.9	1435	5	AAS93062	Aas93062 DNA encod
401	26	92.9	906	6	ABQ15943	Abq15943 Oligonucl	C 474	92.9	1435	5	AAS80861	Aas80861 DNA encod
402	26	92.9	909	10	ADK59700	Adk59700 Plant DNA	475	92.9	1453	6	ABN86418	Abn86418 C. glutam
403	26	92.9	948	5	AAS70518	Aas70518 DNA encod	C 476	92.9	1476	2	AAV72549	AAV72549 Vernonia
404	26	92.9	965	4	ABEL14369	Abel14369 Drosophil	477	92.9	1476	2	AAAF91721	AAAF91721 Porphyrom
405	26	92.9	969	5	ABV14099	Abv14099 Human pro	478	92.9	1506	6	ABN81604	Abn81604 Fungal de
406	26	92.9	981	8	ACA27357	Aca27357 Prokaryot	479	92.9	1506	6	AAAF49927	AAAF49927 Human mol
407	26	92.9	988	5	AAS67559	Aas67559 DNA encod	C 480	92.9	1510	3	AAAF15123	AAAF15123 Trichoder
408	26	92.9	989	3	AAF07624	Aaf07624 Fusarium	481	92.9	1518	5	AAAF15123	AAAF15123 Trichoder
409	26	92.9	993	10	ADH82201	Adh82201 Enterococ	482	92.9	1522	4	ABL29525	AbL29525 Drosophil
410	26	92.9	999	5	AAS66807	Aas66807 DNA encod	483	92.9	1524	8	ACA25780	Aca25780 Prokaryot
411	26	92.9	1001	3	AAC57570	Aac57570 Arachidon	484	92.9	1536	12	ADO61600	Ado61600 Transcrip
412	26	92.9	1024	6	ABX66013	Abx66013 Helicobac	485	92.9	1549	3	AAAL6638	AAAL6638 Human sec
413	26	92.9	1026	4	AAS31357	Aas31357 Human cDN	486	92.9	1552	4	ABL26673	AbL26673 Drosophil
414	26	92.9	1026	6	ABQ66681	Abq66681 Human pol	487	92.9	1560	4	AAF58340	Aaf58340 Human GTP
415	26	92.9	1026	10	ADC10703	Adc10703 Human cDN	488	92.9	1566	6	ABN59803	Abn59803 Novel hum
416	26	92.9	1032	4	AAK84893	Aak84893 Human imm	489	92.9	1573	6	ABN81700	Abn81700 Corynebac
417	26	92.9	1059	6	ABL40697	AbL40697 Human big	C 490	92.9	1580	4	ABL06341	AbL06341 Drosophil
418	26	92.9	1065	12	ABN74382	Abn74382 Thale cre	C 491	92.9	1580	4	ABL20401	AbL20401 Drosophil
419	26	92.9	1074	6	ABN92105	Abn92105 Staphyloc	492	92.9	1602	3	AAC45745	Aac45745 Arabidops
420	26	92.9	1074	8	ACA47522	Aca47522 Prokaryot	493	92.9	1602	6	ABZ13221	Abz13221 Arabidops
421	26	92.9	1074	8	ACA46417	Aca46417 Prokaryot	494	92.9	1602	8	ACA46740	Aca46740 Prokaryot
422	26	92.9	1074	8	ACF74006	Acf74006 Staphyloc	C 495	92.9	1609	10	ADA53937	Ada53937 Human cod
423	26	92.9	1076	4	AAF71884	Aaf71884 Corynebac	496	92.9	1617	5	AAS68206	Aas68206 DNA encod
424	26	92.9	1077	4	AAH52312	Aah52312 S. epider	497	92.9	1620	2	AAAF1594	AAAF1594 Porphyrom
425	26	92.9	1077	5	AAS82659	Aas82659 DNA encod	498	92.9	1620	3	AAC50170	Aac50170 Arabidops
426	26	92.9	1077	8	ACA46196	Aca46196 Prokaryot	499	92.9	1621	5	AAS67924	Aas67924 DNA encod
427	26	92.9	1077	10	ADH48620	Adh48620 Exprimen	500	92.9	1634	10	ADD30984	Add30984 Plant yie
428	26	92.9	1077	10	ADH48618	Adh48618 Branched-	501	92.9	1634	12	ADI44248	Adi44248 Plant tra
429	26	92.9	1080	2	AAT12906	Aat12906 Sardinian	502	92.9	1636	2	AAAF1940	AAAF1940 Arabidops
430	26	92.9	1080	2	AAT12905	Aat12905 Sardinian	503	92.9	1642	2	AAAF1941	Aat191941 Arabidops
431	26	92.9	1080	2	AAT12904	Aat12904 Sardinian	504	92.9	1642	2	AAAF1938	Aat191938 Arabidops
432	26	92.9	1086	3	AAK89448	Aak89448 Branched	505	92.9	1643	2	AAAF1938	Aat191938 Arabidops
433	26	92.9	1089	6	ABN93103	Abn93103 Staphyloc	506	92.9	1663	4	AAS02541	Aas02541 Human sec
434	26	92.9	1109	6	AAV05140	Aav05140 CDNA enco	C 507	92.9	1667	6	ABQ70628	Abq70628 Listeria
435	26	92.9	1109	6	ABQ68493	Abq68493 Listeria	C 508	92.9	1671	4	ABL08483	AbL08483 Drosophil
436	26	92.9	1110	10	ADF02857	Adf02857 Bacterial	C 509	92.9	1671	10	ADB76896	AdB76896 Human N-a
437	26	92.9	1118	11	ABD09897	Abd09897 Pseudomon	C 510	92.9	1671	10	ADE25917	Ade25917 Novel N-a
438	26	92.9	1137	5	AAH67379	Aah67379 C Glutami	511	92.9	1678	4	AAH99230	Aah99230 Human pro
439	26	92.9	1145	2	AAT93288	Aat93287 Tomato ye	512	92.9	1679	4	ABL18073	AbL18073 Drosophil
440	26	92.9	1145	2	AAT93288	Aat93288 Tomato ye	C 513	92.9	1695	2	AAQ84378	Aaq84378 Tomato ye
441	26	92.9	1145	2	AAT93288	Aat93289 Tomato ye	514	92.9	1729	4	ABL22659	AbL22659 Human Na/
442	26	92.9	1145	2	AAT93288	Aat93311 Tomato ye	515	92.9	1729	4	ABL22659	AbL22659 Drosophil
443	26	92.9	1151	3	AAC45524	Aac45524 Arabidops	516	92.9	1734	4	AAAF1142	AaF1142 Human bre
444	26	92.9	1154	3	AAC32906	Aac32906 Arabidops	C 517	92.9	1745	4	AAH16356	Aah16356 Human cDN

c 518	26	92.9	1761	6	ABQ31703	Abq31703 Oligonucle	c 591	26	92.9	2286	4	AAF26586	Aaf26586 DNA encod
c 519	26	92.9	1761	6	ABQ31702	Abq31702 Oligonucle	592	26	92.9	2311	3	AAA26455	Aaa26455 Human sec
c 520	26	92.9	1762	12	ADO35973	Ado35973 Novel mou	593	26	92.9	2311	12	ADL71519	Adl71519 Novel hum
521	26	92.9	1764	4	AD050576	Ad050576 Arabidops	594	26	92.9	2313	3	AAC43117	Aac43117 Arabidops
522	26	92.9	1764	5	AD066661	Ad066661 A. thalia	c 595	26	92.9	2331	6	ABK91320	Abk91320 cDNA enco
523	26	92.9	1764	6	ABZ14708	Abz14708 Arabidops	596	26	92.9	2331	6	ABK91320	Abk91320 cDNA enco
524	26	92.9	1764	12	ADO01802	Ado01802 Thalecres	597	26	92.9	2340	4	AAF71656	Aaf71656 Corynebac
c 525	26	92.9	1768	2	AX336278	Ax336278 Wheat Rht	598	26	92.9	2347	10	ADD13392	Add13392 C. glutam
c 526	26	92.9	1771	12	ADL12299	Adl12299 Human ste	599	26	92.9	2376	5	AAS91752	Aas91752 DNA encod
c 527	26	92.9	1773	6	AB213413	Ab213413 Arabidops	600	26	92.9	2376	8	AAC66688	Aac66688 DNA encod
528	26	92.9	1779	4	AF25480	Af25480 Nucleotid	601	26	92.9	2376	8	ACD05933	Acd05933 Novel hum
c 529	26	92.9	1791	11	ADL22592	Adl22592 Human dis	602	26	92.9	2385	12	ADM47783	Adm47783 Polynucle
c 530	26	92.9	1836	8	ACH24740	Ach24740 Prokaryot	603	26	92.9	2406	11	ADM01554	Adm01554 Human cDN
c 531	26	92.9	1837	2	AAT33852	Aat33852 Rat vas d	604	26	92.9	2422	6	AA037665	Aa037665 Human G-P
532	26	92.9	1917	6	AA149945	Aa149945 Human mol	605	26	92.9	2449	12	ADQ22966	Adq22966 Human sof
c 533	26	92.9	1927	10	AD330876	Ad330876 Plant vie	606	26	92.9	2481	9	ADA09915	Ada09915 Human rec
c 534	26	92.9	1927	10	AD330876	Ad330876 Plant vie	607	26	92.9	2500	12	ADM47776	Adm47776 Polynucle
c 535	26	92.9	1927	12	ADI41890	Adi41890 Plant tra	608	26	92.9	2508	6	ABQ90410	Abq90410 M. capsul
536	26	92.9	1929	6	ABQ69126	Abq69126 Listeria	c 608	26	92.9	2523	10	ACC69598	Acc69598 Rat sweet
537	26	92.9	1946	4	AA112975	Aa112975 Probe #29	c 609	26	92.9	2523	10	ACC69598	Acc69598 Rat sweet
538	26	92.9	1946	4	ABA54677	Ab54677 Human foe	c 610	26	92.9	2529	10	ACC69595	Acc69595 Mouse awe
539	26	92.9	1946	4	ABA54677	Ab54677 Human foe	c 611	26	92.9	2529	12	ADM42818	Adm42818 DNA encod
540	26	92.9	1946	4	ABA44227	Ab44227 Human bre	c 612	26	92.9	2529	12	ADP70063	Adp70063 Mouse TIR
541	26	92.9	1946	4	ABA24459	Ab24459 Probe #29	c 613	26	92.9	2579	3	AZ50744	Az50744 Mouse sen
542	26	92.9	1946	4	ABA28410	Ab28410 Human bon	614	26	92.9	2579	10	ADK90715	Adk90715 Mouse tas
543	26	92.9	1946	4	ABK28006	Abk28006 Human liv	c 616	26	92.9	2601	4	AAK94456	Aak94456 Human ful
544	26	92.9	1946	4	ABK28006	Abk28006 Human liv	c 617	26	92.9	2601	12	ADL31227	Adl31227 Full leng
545	26	92.9	1946	5	AAI02893	Aai02893 Probe #28	c 618	26	92.9	2610	12	ADL31227	Adl31227 Full leng
546	26	92.9	1946	6	ABK02918	Abk02918 Human gen	619	26	92.9	2636	5	ABA09697	Ab09697 Human bon
547	26	92.9	1948	4	AAH17684	Aah17684 Human cDN	c 620	26	92.9	2636	5	ABA09697	Ab09697 Human bon
548	26	92.9	1951	4	AA005791	Aa005791 Arabidops	c 621	26	92.9	2651	10	ADA53171	Ada53171 Human cod
549	26	92.9	1951	5	AA006646	Aa006646 A. thalia	622	26	92.9	2711	4	ABL03426	Ab103426 Drosophil
550	26	92.9	1951	10	ADD55687	Ad55687 Thalecres	c 623	26	92.9	2711	4	ABL03426	Ab103426 Drosophil
551	26	92.9	1951	10	ADD30751	Ad30751 Plant vie	c 624	26	92.9	2771	3	AZ50743	Az50743 Rat senso
552	26	92.9	1951	12	ADI43826	Adi43826 Plant tra	c 625	26	92.9	2771	10	ADK90714	Adk90714 Rat taste
553	26	92.9	1951	12	ADO01804	Ado01804 Thalecres	626	26	92.9	2813	2	AA113396	Aa113396 Enterococ
c 554	26	92.9	1956	10	ADC10115	Adc10115 Human NOV	c 627	26	92.9	2813	2	AA113396	Aa113396 Enterococ
555	26	92.9	1961	4	ABL04017	Ab104017 Drosophil	628	26	92.9	2818	12	ADH22396	Adh22396 Human cDN
556	26	92.9	1964	2	AAAT91937	Aat91937 Arabidops	c 629	26	92.9	2831	2	AAV52942	Aav52942 Rat UNC-5
c 557	26	92.9	1969	8	ACC46456	Acc46456 Human dit	630	26	92.9	2838	12	ADO09501	Ado09501 Rat trans
558	26	92.9	1975	6	ABK35064	Abk35064 Human cDN	c 631	26	92.9	2860	6	ABT06279	Abt06279 Human NOV
559	26	92.9	2000	10	ACC61575	Acc61575 Gene sequ	c 632	26	92.9	2860	6	ABT06280	Abt06280 Human NOV
560	26	92.9	2000	10	ADK64425	Adk64425 Disease t	c 633	26	92.9	2885	2	AXA40055	Axa40055 Colon can
561	26	92.9	2008	6	ABQ70601	Abq70601 Listeria	634	26	92.9	2885	10	AAD54023	Aad54023 Human col
562	26	92.9	2008	6	ABQ68952	Abq68952 Listeria	c 635	26	92.9	2895	6	ABQ93898	Abq93898 Human tra
563	26	92.9	2038	2	AAQ40490	Aaq40490 Clone FG5	636	26	92.9	2895	12	ADO10076	Ado10076 Novel hum
564	26	92.9	2091	10	ADB63519	Abd63519 Human cDN	c 637	26	92.9	2915	10	ADG62879	Adg62879 Human cDN
565	26	92.9	2106	8	ABK62890	Abk62890 Human act	638	26	92.9	2975	12	ADO35871	Ado35871 Novel mou
566	26	92.9	2125	2	AX336279	Ax336279 Wheat Rht	c 639	26	92.9	2995	6	ABK92062	Abk92062 DNA encod
567	26	92.9	2133	11	ABD09771	Abd09771 Pseudomon	c 640	26	92.9	3000	6	ABQ70742	Abq70742 Listeria
568	26	92.9	2159	4	AAF71657	Aaf71657 Corynebac	c 641	26	92.9	3001	3	AAH51737	Aah51737 Chromosom
c 569	26	92.9	2162	4	ABL05685	Ab105685 Drosophil	642	26	92.9	3021	4	ABL23671	Ab123671 Drosophil
570	26	92.9	2166	12	ADQ23221	Adq23221 Human sof	643	26	92.9	3024	2	AAT05501	Aat05501 Bfrr-PvuII
571	26	92.9	2184	3	ACA43013	Acc43013 Arabidops	644	26	92.9	3028	4	ABL14368	Ab114368 Drosophil
572	26	92.9	2184	6	ABZ14856	Abz14856 Arabidops	645	26	92.9	3072	6	ABN67056	Abn67056 Streptoco
573	26	92.9	2202	2	AAQ80345	Aaq80345 DNA fragm	c 646	26	92.9	3076	3	AAZ24477	Aaz24477 H. viresc
574	26	92.9	2202	2	AAV92904	Aav92904 Corynefor	647	26	92.9	3109	3	ABS69893	Abs69893 Human rec
575	26	92.9	2203	2	AAV15031	Aav15031 Promoter	648	26	92.9	3155	8	ACA50444	Aca50444 Prokaryot
576	26	92.9	2213	4	AAK94352	Aak94352 Human ful	649	26	92.9	3177	8	ACA48237	Acc48237 Prokaryot
577	26	92.9	2213	12	ADL31028	Adl31028 Full leng	650	26	92.9	3177	8	ACA48237	Acc48237 Prokaryot
578	26	92.9	2217	5	AAH68453	Aah68453 C glutami	651	26	92.9	3180	8	ACA33021	Acca33021 Prokaryot
579	26	92.9	2225	11	ADN95102	Adn95102 Human LEC	652	26	92.9	3183	8	ACA33598	Acca33598 Prokaryot
c 580	26	92.9	2233	4	ABK89558	Abk89558 Human his	653	26	92.9	3201	5	AAH77680	Aah77680 DNA encod
c 581	26	92.9	2233	6	ABK87719	Abk87719 Human cDN	654	26	92.9	3204	4	AAH54164	Aah54164 S. epider
c 582	26	92.9	2233	10	ADC21696	Adc21696 Human cDN	655	26	92.9	3225	10	ADH84800	Adh84800 Enterococ
583	26	92.9	2240	3	ACA26419	Aca26419 Human sec	656	26	92.9	3245	12	ADO35451	Ado35451 Novel mou
584	26	92.9	2240	12	ADL71480	Adl71480 Novel hum	657	26	92.9	3285	12	ADJ34927	Adj34927 DNA encod
585	26	92.9	2244	10	ADE63224	Ade63224 Human gen	658	26	92.9	3300	6	ABS55635	Abs55635 Mouse sco
586	26	92.9	2244	10	ADD46320	Add46320 Human gen	659	26	92.9	3308	12	ADN00736	Adn00736 Human LDL
587	26	92.9	2255	2	AA336280	Aa336280 Maize lal	660	26	92.9	3314	2	ADG62711	Adg62711 Human cDN
c 588	26	92.9	2258	8	ADA89812	Ada89812 Staphyloc	661	26	92.9	3330	2	AAQ88687	Aaq88687 Human ver
589	26	92.9	2264	5	ADG20698	Adg20698 Human ABC	c 662	26	92.9	3341	4	AAH29722	Aah29722 S cerevis
590	26	92.9	2283	3	AAC50691	Aac50691 Arabidops	c 663	26	92.9	3359	12	ADH43337	Adh43337 Human cla

c 664	26	92.9	3400	10	ADJ95167	Adj95167 Novel NOV	737	26	92.9	3884	9	ADA76750	Ada76750 Human PRO
c 665	26	92.9	3400	12	ADH41652	Adh41652 Novel hum	738	26	92.9	3884	9	ADA89380	Ada89380 Novel hum
c 666	26	92.9	3447	3	AAC50274	Aac50274 Arabidops	739	26	92.9	3884	9	ADA97385	Ada97385 Human PRO
c 667	26	92.9	3449	10	ADK67008	Adk67008 Gene #98	740	26	92.9	3884	9	ADB27142	Adb27142 cDNA enco
c 668	26	92.9	3501	5	ADG20622	Adg20622 Human ABC	741	26	92.9	3884	9	ADB22075	Adb22075 Novel hum
c 669	26	92.9	3501	11	ADN39805	Adn39805 Cancer/an	742	26	92.9	3884	9	ADA66766	Ada66766 Human PRO
c 670	26	92.9	3531	8	ABX62888	Abx62888 Human act	743	26	92.9	3884	9	ADB22627	Adb22627 Human PRO
c 671	26	92.9	3531	12	ADL13354	Adl13354 Human ste	744	26	92.9	3884	9	ADB23400	Adb23400 Human PRO
c 672	26	92.9	3571	6	ABZ11529	Abz11529 Human pol	745	26	92.9	3884	9	ADA92122	Ada92122 Novel hum
c 673	26	92.9	3571	12	ADM44047	Adm44047 Novel hum	746	26	92.9	3884	9	ADB15185	Adb15185 Human PRO
c 674	26	92.9	3604	4	AAH54835	Aah54835 S. epider	747	26	92.9	3884	9	ADB38437	Adb38437 Novel hum
c 675	26	92.9	3622	8	ABX62889	Abx62889 Human act	748	26	92.9	3884	9	ADB37885	Adb37885 Novel hum
c 676	26	92.9	3656	2	RAT36751	Rat36751 VLDL rece	749	26	92.9	3884	9	ADB66357	Adb66357 Novel hum
c 677	26	92.9	3656	12	ADQ22405	Adq22405 Human sof	750	26	92.9	3884	10	ADB89437	Adb89437 Human PRO
c 678	26	92.9	3680	4	ABL26672	AbL26672 Drosophil	751	26	92.9	3884	10	ADB90169	Adb90169 Human PRO
c 679	26	92.9	3705	12	ADJ35089	Adj35089 DNA enco	752	26	92.9	3884	10	ADB39270	Adb39270 Novel hum
c 680	26	92.9	3865	4	ABL06482	AbL06482 Drosophil	753	26	92.9	3884	10	ADB46893	Adb46893 Novel hum
c 681	26	92.9	3884	4	AA521316	Aa521316 Human cDN	754	26	92.9	3884	10	ADB46500	Adb46500 Human PRO
c 682	26	92.9	3884	8	ACA03675	Aca03675 cDNA enco	755	26	92.9	3884	10	ADB77105	Adb77105 Novel hum
c 683	26	92.9	3884	8	ACX89213	Acx89213 DNA enco	756	26	92.9	3884	10	ADB34262	Adb34262 Human PRO
c 684	26	92.9	3884	8	ACD41867	Acd41867 Human sec	757	26	92.9	3884	10	ADB35366	Adb35366 Human PRO
c 685	26	92.9	3884	8	ACA04096	Aca04096 Human cDN	758	26	92.9	3884	10	ADB33710	Adb33710 Human PRO
c 686	26	92.9	3884	8	ADA45664	Ada45664 Novel hum	759	26	92.9	3884	10	ADB34814	Adb34814 Human PRO
c 687	26	92.9	3884	9	ADA76095	Ada76095 Human PRO	760	26	92.9	3884	10	ADB35918	Adb35918 Human PRO
c 688	26	92.9	3884	9	ADA18745	Ada18745 Human PRO	761	26	92.9	3884	10	ADB46313	Adb46313 Novel hum
c 689	26	92.9	3884	9	ADA61368	Ada61368 Homo sapi	762	26	92.9	3884	10	ADC50186	Adc50186 Novel hum
c 690	26	92.9	3884	9	ADB19153	Adb19153 Novel hum	763	26	92.9	3884	10	ADC71733	Adc71733 Novel hum
c 691	26	92.9	3884	9	ADB27694	Adb27694 cDNA enco	764	26	92.9	3884	10	ADC59712	Adc59712 Novel hum
c 692	26	92.9	3884	9	ADA85621	Ada85621 Novel hum	765	26	92.9	3884	10	ADC52719	Adc52719 Novel hum
c 693	26	92.9	3884	9	ADA86173	Ada86173 Novel hum	766	26	92.9	3884	10	ADC57073	Adc57073 Novel hum
c 694	26	92.9	3884	9	ADA815737	Ada815737 Human PRO	767	26	92.9	3884	10	ADC50739	Adc50739 Novel hum
c 695	26	92.9	3884	9	ADA47523	Ada47523 Human PRO	768	26	92.9	3884	10	ADC60264	Adc60264 Novel hum
c 696	26	92.9	3884	9	ADB30325	Adb30325 cDNA enco	769	26	92.9	3884	10	ADC50739	Adc50739 Novel hum
c 697	26	92.9	3884	9	ADA85621	Ada85621 Novel hum	770	26	92.9	3884	10	ADC54364	Adc54364 Novel hum
c 698	26	92.9	3884	9	ADA96833	Ada96833 Human PRO	771	26	92.9	3884	10	ADC53325	Adc53325 Novel hum
c 699	26	92.9	3884	9	ADA79137	Ada79137 Human PRO	772	26	92.9	3884	10	ADC58848	Adc58848 Novel hum
c 700	26	92.9	3884	9	ADA87276	Ada87276 Novel hum	773	26	92.9	3884	10	ADC55726	Adc55726 Novel hum
c 701	26	92.9	3884	9	ADB16478	Adb16478 Human PRO	774	26	92.9	3884	10	ADC58296	Adc58296 Novel hum
c 702	26	92.9	3884	9	ADA91570	Ada91570 Novel hum	775	26	92.9	3884	10	ADD02970	Add02970 Novel hum
c 703	26	92.9	3884	9	ADB14633	Adb14633 Human PRO	776	26	92.9	3884	10	ADC89962	Adc89962 Novel hum
c 704	26	92.9	3884	9	ADB18594	Adb18594 Novel hum	777	26	92.9	3884	10	ADC69381	Adc69381 cDNA enco
c 705	26	92.9	3884	9	ADA93809	Ada93809 Human PRO	778	26	92.9	3884	10	ADC48270	Adc48270 Human PRO
c 706	26	92.9	3884	9	ADB19705	Adb19705 Novel hum	779	26	92.9	3884	10	ADD09799	Add09799 Human PRO
c 707	26	92.9	3884	9	ADB13017	Adb13017 Human PRO	780	26	92.9	3884	10	ADD04374	Add04374 Novel hum
c 708	26	92.9	3884	9	ACD98496	Acd98496 Novel hum	781	26	92.9	3884	10	ADC80330	Adc80330 Novel hum
c 709	26	92.9	3884	9	ADA74271	Ada74271 Human PRO	782	26	92.9	3884	10	ADD10837	Add10837 Human PRO
c 710	26	92.9	3884	9	ADB24504	Adb24504 Human PRO	783	26	92.9	3884	10	ADC47718	Adc47718 Human PRO
c 711	26	92.9	3884	9	ADA82028	Ada82028 Human PRO	784	26	92.9	3884	10	ADC79778	Adc79778 Novel hum
c 712	26	92.9	3884	9	ADA74991	Ada74991 Human PRO	785	26	92.9	3884	10	ADD09247	Add09247 Human PRO
c 713	26	92.9	3884	9	ADA85069	Ada85069 Novel hum	786	26	92.9	3884	10	ADD40960	Add40960 Novel hum
c 714	26	92.9	3884	9	ADA84517	Ada84517 Novel hum	787	26	92.9	3884	10	ADD52099	Add52099 cDNA enco
c 715	26	92.9	3884	9	ADB29773	Adb29773 cDNA enco	788	26	92.9	3884	10	ADD52839	Add52839 cDNA enco
c 716	26	92.9	3884	9	ADA80301	Ada80301 Human PRO	789	26	92.9	3884	10	ADD53391	Add53391 Novel hum
c 717	26	92.9	3884	9	ADA75543	Ada75543 Human PRO	790	26	92.9	3884	10	ADD51547	Add51547 cDNA enco
c 718	26	92.9	3884	9	ADA46768	Ada46768 Human PRO	791	26	92.9	3884	10	ADD02346	Add02346 Human PRO
c 719	26	92.9	3884	9	ADB25064	Adb25064 Human PRO	792	26	92.9	3884	10	ADD001780	Add001780 Novel hum
c 720	26	92.9	3884	9	ADA93240	Ada93240 Human PRO	793	26	92.9	3884	10	ADD53962	Add53962 Novel hum
c 721	26	92.9	3884	9	ADB26590	Adb26590 cDNA enco	794	26	92.9	3884	10	ADD92279	Add92279 Human PRO
c 722	26	92.9	3884	9	ADB30877	Adb30877 cDNA enco	795	26	92.9	3884	10	ADD91175	Add91175 Human PRO
c 723	26	92.9	3884	9	ADA60805	Ada60805 Homo sapi	796	26	92.9	3884	10	ADB32086	Adb32086 Novel hum
c 724	26	92.9	3884	9	ADB23952	Adb23952 Human PRO	797	26	92.9	3884	10	ADB32086	Adb32086 Novel hum
c 725	26	92.9	3884	9	ADA96281	Ada96281 Human PRO	798	26	92.9	3884	10	ADB22018	Adb22018 cDNA enco
c 726	26	92.9	3884	9	ADA80853	Ada80853 Human PRO	799	26	92.9	3884	10	ADD79242	Add79242 cDNA enco
c 727	26	92.9	3884	9	ADA95729	Ada95729 Human PRO	800	26	92.9	3884	10	ADE41778	Ade41778 Human PRO
c 728	26	92.9	3884	9	ADB26038	Adb26038 cDNA enco	801	26	92.9	3884	10	ADE17595	Ade17595 Human PRO
c 729	26	92.9	3884	9	ADB21523	Adb21523 Novel hum	802	26	92.9	3884	10	ADD91727	Add91727 Human PRO
c 730	26	92.9	3884	9	ADA77302	Ada77302 Human PRO	803	26	92.9	3884	10	ADB33190	Adb33190 Novel hum
c 731	26	92.9	3884	9	ADB18042	Adb18042 cDNA enco	804	26	92.9	3884	10	ADB33742	Adb33742 Novel hum
c 732	26	92.9	3884	9	ADA86725	Ada86725 Novel hum	805	26	92.9	3884	10	ADD79794	Add79794 cDNA enco
c 733	26	92.9	3884	9	ADA87828	Ada87828 Novel hum	806	26	92.9	3884	10	ADD92831	Add92831 Human PRO
c 734	26	92.9	3884	9	ADA46216	Ada46216 Novel hum	807	26	92.9	3884	10	ADE19251	Ade19251 Human PRO
c 735	26	92.9	3884	9	ADB28246	Adb28246 cDNA enco	808	26	92.9	3884	10	ADE18699	Ade18699 Human PRO
c 736	26	92.9	3884	9	ADB28798	Adb28798 cDNA enco	809	26	92.9	3884	10	ADE42895	Ade42895 Human PRO

810	26	92.9	3884	10	ADD95684	Ad955684 Human PRO	883	26	92.9	3884	12	ADG57230	Novel hum
811	26	92.9	3884	10	ADE22570	Ade22570 cDNA enco	884	26	92.9	3884	12	ADG56678	Novel hum
812	26	92.9	3884	10	ADD78688	Ad78688 cDNA enco	885	26	92.9	3884	12	ADG55574	Novel hum
813	26	92.9	3884	10	ADE32638	Ade32638 Novel hum	886	26	92.9	3884	12	ADG58334	Novel hum
814	26	92.9	3884	10	ADE42330	Ade42330 Human PRO	887	26	92.9	3884	12	ADG70700	Novel hum
815	26	92.9	3884	10	ADD80346	Ad80346 cDNA enco	888	26	92.9	3884	12	ADG57782	Novel hum
816	26	92.9	3884	10	ADD89374	Ad89374 Human PRO	889	26	92.9	3884	12	ADG53366	Novel hum
817	26	92.9	3884	10	ADE40658	Ade40658 Human PRO	890	26	92.9	3884	12	ADG71252	Novel hum
818	26	92.9	3884	10	ADE04457	Ade04457 Human PRO	891	26	92.9	3884	12	ADG81439	Human PRO
819	26	92.9	3884	10	ADE92586	Ade92586 Human PRO	892	26	92.9	3884	12	ADH30401	Human PRO
820	26	92.9	3884	10	ADG21295	Adg21295 Novel hum	893	26	92.9	3884	12	ADH11768	Novel hum
821	26	92.9	3884	10	ADG22936	Adg22936 Novel hum	894	26	92.9	3884	12	ADG52190	Novel hum
822	26	92.9	3884	10	ADF97271	Adf97271 Human PRO	895	26	92.9	3884	12	ADG53918	Novel hum
823	26	92.9	3884	10	ADG80335	Adg80335 Human PRO	896	26	92.9	3884	12	ADG80887	Human PRO
824	26	92.9	3884	10	ADG79783	Adg79783 Human PRO	897	26	92.9	3884	12	ADG56126	Novel hum
825	26	92.9	3884	10	ADH55075	Adh55075 Novel hum	898	26	92.9	3884	12	ADH12392	Novel hum
826	26	92.9	3884	10	ADH55627	Adh55627 Novel hum	899	26	92.9	3884	12	ADG61238	Novel hum
827	26	92.9	3884	10	ADI64795	Adi64795 Novel hum	900	26	92.9	3884	12	ADH28325	Human PRO
828	26	92.9	3884	10	ADI63294	Adi63294 Novel hum	901	26	92.9	3884	12	ADG54470	Novel hum
829	26	92.9	3884	10	ADH81708	Adh81708 Novel hum	902	26	92.9	3884	12	ADG59510	Novel hum
830	26	92.9	3884	10	ADH81156	Adh81156 Novel hum	903	26	92.9	3884	12	ADI80934	cDNA enco
831	26	92.9	3884	10	ACD23925	Adc23925 Novel hum	904	26	92.9	3884	12	ADG09677	Novel hum
832	26	92.9	3884	10	ACA67066	Aca67066 cDNA enco	905	26	92.9	3884	12	ADI15148	Novel hum
833	26	92.9	3884	11	ADM82325	Adm82325 Novel hum	906	26	92.9	3884	12	ADG09025	Novel hum
834	26	92.9	3884	11	ADN15724	Adn15724 Novel hum	907	26	92.9	3884	12	ADI14480	Novel hum
835	26	92.9	3884	11	ADN16353	Adn16353 Novel hum	908	26	92.9	3884	12	ADI18075	Novel hum
836	26	92.9	3884	11	ADN15172	Adn15172 Novel hum	909	26	92.9	3884	12	ADJ63356	Novel hum
837	26	92.9	3884	11	ADN14620	Adn14620 Novel hum	910	26	92.9	3884	12	ADJ77251	Human PRO
838	26	92.9	3884	12	ADC80882	Adc80882 Novel hum	911	26	92.9	3884	12	ADJ65373	cDNA enco
839	26	92.9	3884	12	ADD76330	Ad76330 Human PRO	912	26	92.9	3884	12	ADM27509	cDNA enco
840	26	92.9	3884	12	ADD87694	Ad87694 Human PRO	913	26	92.9	3884	12	ADM42233	cDNA enco
841	26	92.9	3884	12	ADH86098	Adh86098 Human PRO	914	26	92.9	3884	12	ADM28095	cDNA enco
842	26	92.9	3884	12	ADH75546	Adh75546 Human PRO	915	26	92.9	3966	8	ABX77586	Different
843	26	92.9	3884	12	ADE23122	Ade23122 cDNA enco	916	26	92.9	3966	12	ADL12826	Human ste
844	26	92.9	3884	12	ADE23674	Ade23674 cDNA enco	917	26	92.9	4004	4	AAH18531	Human cDN
845	26	92.9	3884	12	ADE24317	Ade24317 cDNA enco	918	26	92.9	4021	3	AACT77217	Human ORF
846	26	92.9	3884	12	ADD87142	Ad87142 Human PRO	919	26	92.9	4090	4	AAH54270	S. epider
847	26	92.9	3884	12	ADE89008	Ade89008 Human PRO	920	26	92.9	4140	4	ABL04016	Drosophil
848	26	92.9	3884	12	ADE18147	Ade18147 Human PRO	921	26	92.9	4146	6	ABZ12888	Arabidops
849	26	92.9	3884	12	ADE88456	Ade88456 Human PRO	922	26	92.9	4146	12	ADM57192	A thalian
850	26	92.9	3884	12	ADE94476	Ade94476 cDNA enco	923	26	92.9	4304	4	ABL06340	Drosophil
851	26	92.9	3884	12	ADE90887	Ade90887 Human PRO	924	26	92.9	4338	3	AACT6239	Arabidops
852	26	92.9	3884	12	ADE95028	Ade95028 cDNA enco	925	26	92.9	4380	12	ADP18652	Human TAT
853	26	92.9	3884	12	ADE93138	Ade93138 Human PRO	926	26	92.9	4381	8	ACC50192	Breast ca
854	26	92.9	3884	12	ADF34719	Adf34719 cDNA enco	927	26	92.9	4538	2	ABL11482	Drosophil
855	26	92.9	3884	12	ADG92034	Adg92034 Novel hum	928	26	92.9	4578	2	AAV64591	Tobacco c
856	26	92.9	3884	12	ADE90335	Ade90335 Human PRO	929	26	92.9	4854	4	ABL22658	Drosophil
857	26	92.9	3884	12	ADE91482	Ade91482 Novel hum	930	26	92.9	4972	4	ABL20400	Drosophil
858	26	92.9	3884	12	ADG02061	Adg02061 Human PRO	931	26	92.9	5193	3	AACT76457	Human ORF
859	26	92.9	3884	12	ADG21847	Adg21847 Novel hum	932	26	92.9	5352	4	AA504049	Human ABC
860	26	92.9	3884	12	ADG19917	Adg19917 cDNA enco	933	26	92.9	5352	6	ABL58166	Human ABC
861	26	92.9	3884	12	ADF97823	Adf97823 Human PRO	934	26	92.9	5355	4	AAK52263	Human pol
862	26	92.9	3884	12	ADG24040	Adg24040 Novel hum	935	26	92.9	5367	6	ABL34545	Human met
863	26	92.9	3884	12	ADF98394	Adf98394 Human PRO	936	26	92.9	5367	6	ABN80153	Human che
864	26	92.9	3884	12	ADG03225	Adg03225 Human PRO	937	26	92.9	5491	4	AA546390	Tumour su
865	26	92.9	3884	12	ADG98946	Adg98946 Human PRO	938	26	92.9	5520	10	ADE53740	Human pro
866	26	92.9	3884	12	ADG16531	Adg16531 cDNA enco	939	26	92.9	5600	4	AAI60415	Human pol
867	26	92.9	3884	12	ADG04990	Adg04990 Human PRO	940	26	92.9	5816	6	ABK40012	Human che
868	26	92.9	3884	12	ADG19257	Adg19257 cDNA enco	941	26	92.9	5845	6	ABL32300	Human imm
869	26	92.9	3884	12	ADG23488	Adg23488 Novel hum	942	26	92.9	5857	6	ABL32528	Human imm
870	26	92.9	3884	12	ADG03777	Adg03777 Human PRO	943	26	92.9	5930	4	AAI36019	Human mus
871	26	92.9	3884	12	ADG08151	Adg08151 Novel hum	944	26	92.9	5930	8	ABX59007	cDNA enco
872	26	92.9	3884	12	ADG15321	Adg15321 cDNA enco	945	26	92.9	5930	12	ADJ29757	Human mus
873	26	92.9	3884	12	ADG96719	Adg96719 Human PRO	946	26	92.9	6002	5	AA584616	DNA enco
874	26	92.9	3884	12	ADG05904	Adg05904 Human PRO	947	26	92.9	6027	4	AAI58629	Human pol
875	26	92.9	3884	12	ADG23488	Adg23488 Novel hum	948	26	92.9	6027	5	ADQ98847	DNA enco
876	26	92.9	3884	12	ADG03777	Adg03777 Human PRO	949	26	92.9	6027	9	ADB48607	Novel hum
877	26	92.9	3884	12	ADG24678	Adg24678 Novel hum	950	26	92.9	6291	4	AA546452	Tumour su
878	26	92.9	3884	12	ADG06975	Adg06975 Novel hum	951	26	92.9	6433	2	AAI13041	Enterococ
879	26	92.9	3884	12	ADG07527	Adg07527 Novel hum	952	26	92.9	6433	6	ABS98836	Enterococ
880	26	92.9	3884	12	ADG55022	Adg55022 Novel hum	953	26	92.9	6439	10	ADB78971	Tick chit
881	26	92.9	3884	12	ADG60686	Adg60686 Novel hum	954	26	92.9	6444	2	AAV74544	Staphyloc
882	26	92.9	3884	12	ADG61790	Adg61790 Novel hum	955	26	92.9	6453	6	ABQ67164	Human ang

956 26 92.9 6517 4 ABL18072 Drosophil
c 957 26 92.9 7713 4 ABL23670 Drosophil
958 26 92.9 8201 4 ABL32307 Human imm
959 26 92.9 8201 6 ABL54328 Chemicall
960 26 92.9 8526 5 ABL18356 Human ner
c 961 26 92.9 8551 5 ABL16478 Human ner
c 962 26 92.9 8551 5 ABL16477 Human ner
963 26 92.9 9302 4 AAS31567 Human DNA
964 26 92.9 9302 6 ABO66891 Human pol
965 26 92.9 9302 10 ADC11178 Human DNA
966 26 92.9 9592 2 AAT36752 Adenoviru
967 26 92.9 9652 6 ABL32909 Human imm
968 26 92.9 10048 6 ABQ67016 Human ang
969 26 92.9 10138 6 ABK28344 DNA trans
970 26 92.9 10860 4 ABL06548 Drosophil
971 26 92.9 11270 4 ABL110066 Drosophil
972 26 92.9 11334 6 ABL33232 Human imm
973 26 92.9 11519 10 ADE09753 Novel DNA
974 26 92.9 11520 4 AAH98291 Human EST
975 26 92.9 11522 6 ABK89709 Oestrogen
976 26 92.9 12012 4 ABL29524 Drosophil
c 977 26 92.9 12033 4 ABL09407 Drosophil
978 26 92.9 12142 6 ABL33672 Human imm
979 26 92.9 12712 12 ADN17332 Cryphonoc
980 26 92.9 12752 2 AAQ47380 Sequence
981 26 92.9 13560 4 ABL28568 Drosophil
982 26 92.9 13862 4 AAL02789 Human rep
983 26 92.9 13862 4 AAL07516 Human rep
984 26 92.9 13862 4 ABA08208 Human imm
985 26 92.9 16086 4 AAK69721 Human mus
c 986 26 92.9 16086 4 AAL36020 Human mus
c 987 26 92.9 16086 8 ABX59008 cDNA enco
c 988 26 92.9 16086 12 ADJ29758 Human mus
989 26 92.9 17294 6 ABL32987 Human imm
990 26 92.9 19087 6 ABL32792 Human imm
991 26 92.9 19619 2 ABK47192 Mouse Con
c 992 26 92.9 20072 2 AAX13026 Enterococ
c 993 26 92.9 20072 6 ABS98821 Enterococ
994 26 92.9 22055 9 ADA02879 Mouse Pap
995 26 92.9 22055 10 ADB72617 Mouse Pap
996 26 92.9 22055 10 ADC85358 Mouse Dpt
997 26 92.9 22055 12 ADM74474 Murine ca
998 26 92.9 24964 6 ABK86595 Human SA
999 26 92.9 25467 12 ADQ07566 Streptomy
c1000 26 92.9 25891 4 ABL05684 Drosophil

ALIGNMENTS

RESULT 1
ID ABLN76795 standard; cDNA; 318 BP.
XX ABLN76795;
AC ABLN76795;
XX ABLN76795;
DT 08-JUL-2002 (first entry)
XX ABLN76795;
DE Human reductase-like ORF1742 cDNA, SEQ ID NO:3483.

Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
disease monitoring; cytokine; cell proliferation; cell differentiation;
immune modulation; haematopoiesis regulation; tissue growth;
angiogenesis; activin; inhibition; chemotactic; chemokinetic; haemostatic;
thrombolytic; tumour inhibition; bodily characteristic; fertility;
behaviour; cancer; proliferative disorder; neurological disorder;
cardiovascular disease; immune system disorder; organ transplantation;
tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
hypothyroidism; cholesterol ester storage disease; infection; vulvury;
vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
dermatological; analgesic; virucide; antibacterial; fungicide; gene; ss.

OS Homo sapiens.
XX WO200190366-A2.
XX 29-NOV-2001.
XX 24-MAY-2001; 2001WO-US017076.
XX 24-MAY-2000; 2000US-0206690P.
XX (CURA-) CURAGEN CORP.
XX Leach MD, Shinkets RA;
XX WPI; 2002-106200/14.
XX P-PSDB; ABP32769.
XX Novel human polypeptides and polynucleotides useful for diagnosing,
preventing and treating cardiovascular disease, neurodegenerative,
hyperproliferative disorders and disorders related to organ
transplantation.

Claim 1; Page 1131; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
designated ORF (open reading frame) 1-4534, and sequences ABN75054-
ABN79587 represent cDNAs encoding them. The invention also encompasses
polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
referred to as ORFX) proteins, polynucleotides at least 85% identical to
the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
polynucleotides, the recombinant production of ORFX proteins, antibodies
specific for ORFX proteins, methods of detecting ORFX polynucleotides and
polypeptides, and methods of screening for modulators of ORFX expression or
activity, and methods of screening individuals for a predisposition to an
ORFX-associated disorder. The ORFX proteins of the invention have a wide
range of biological activities, such as cytokine, cell proliferation,
cell differentiation, immune modulation, haematopoiesis regulation,
tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
chemokinetic activity, haemostatic activity, thrombolytic activity,
receptor/ligand, antiinflammatory activity, tumour inhibition activity,
and antiinfective activity, and may also be involved in the determination
of bodily characteristics, fertility and behaviour. ORFX proteins,
nucleic acids and antibodies may be used in the treatment of cancers,
other proliferative disorders such as psoriasis and benign tumours,
neurological disorders such as epilepsy and Alzheimer's disease,
cardiovascular diseases, immune system disorders, disorders related to
organ transplantation, disorders of tissue growth and regeneration,
diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
storage disease, and infectious diseases caused by viral, bacterial,
fungal and other pathogens. ORFX nucleic acids may also be used as a
source of primers and probes, in the detection of ORFX genomic sequences
or transcripts, in the identification and cloning of homologous
sequences, in genetic diagnosis, and in forensic biology. The ORFX
nucleic acids may additionally be used to produce transgenic animals
which may be useful for studying the function and/or activity of ORFX
protein, and in drug screening. The ORFX proteins may also be used as
immunogens to generate specific antibodies, which are useful in the
diagnosis, treatment and monitoring of ORFX-associated diseases

XX Sequence 318 BP; 81 A; 71 C; 83 G; 83 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 331 Length: 318
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x ABLN76795 (1-318)

Oy 1 GlyTyr***ValGluGlu 6
||||| |||||||

Db 165 GGATATTCTGTTGAAGAA 182

RESULT 2

AAQ04657

ID AAQ04657 standard; DNA; 330 BP.

XX

AC AAQ04657;

XX

DT 01-OCT-1990 (first entry)

XX

DE Sequence encoding metal chelate specific light chain antigen.

XX

KW Chelate; metal specific; tumour; cancer; ds.

XX

OS Synthetic.

XX

PN EP369567-A.

XX

PD 23-MAY-1990.

XX

PF 08-MAR-1989; 89EP-00302314.

XX

PR 17-NOV-1988; 88US-00274106.

XX

PA (HYBR-) HYBRITECH INC.

XX

PI Johnson MJ;

XX

DR WPI; 1990-157696/21.

DR

DR P-FSDB; AAR05038.

XX

PT Chimeric antibodies directed against metal chelates - obtd. using DNA

PT constructs for chelate-specific light and heavy chain variable regions.

XX

PS Claim 2; Page 28-29; 38pp; English.

XX

CC MAB Variable sequence may be coupled with the constant region from a

CC different mammal to produce a chimeric antibody recognising the specific

CC chelate ion. This may be useful in encouraging tumours to take up

CC radioactivity

XX

SQ Sequence 330 BP; 87 A; 85 C; 80 G; 78 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	345	Length:	330
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	2	Gaps:	0

US-10-030-194A-5 (1-6) x AAQ04657 (1-330)

Qy 1 GlyTyr***ValGluGlu 6

Db 293 GGGTATTCTGTTGAAGAA 310

RESULT 3

AAQ04594

ID AAQ04594 standard; DNA; 330 BP.

XX

AC AAQ04594;

XX

DT 28-SEP-1990 (first entry)

XX

DE Sequence encoding region of chelate-specific light chain antigen.

XX

KW Carcinoembryonic antigen; CEA; Chelate-specific antigen; CHA; tumor;

KW cancer; serum sickness.

XX

OS Homo sapiens.

XX

PN EP369566-A.

XX

PD 23-MAY-1990.

XX

PF 08-MAR-1989; 89EP-00302313.

XX

PR 17-NOV-1988; 88US-00274105.

XX

PA (HYBR-) HYBRITECH INC.

XX

PI Johnson MJ, Phelps JL;

XX

DR WPI; 1990-157695/21.

DR

DR P-FSDB; AAR04938.

XX

PT Bifunctional chimeric antibodies - having variable regions which

PT recognise different antigens and metal chelates and human constant

PT regions.

XX

PS Claim 10; Page 25; 40pp; English.

XX

CC Gene encodes portion of chelate-specific light chain antigen, useful in

CC generating chimeric monoclonal antibody binding to CEA at a tumour site

CC and a metal chelate bonded to say a toxin or other drug

XX

SQ Sequence 330 BP; 86 A; 86 C; 80 G; 78 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	345	Length:	330
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	2	Gaps:	0

US-10-030-194A-5 (1-6) x AAQ04594 (1-330)

Qy 1 GlyTyr***ValGluGlu 6

Db 293 GGGTATTCTGTTGAAGAA 310

RESULT 4

AAQ68744

ID AAQ68744 standard; DNA; 330 BP.

XX

AC AAQ68744;

XX

DT 19-FEB-1995 (first entry)

XX

DE CHA255 light chain variable region coding region.

XX

KW Polymerase chain reaction; primer; PCR; amplify; heavy; light; chain;

KW complementarity determining region; CDR; variable; constant; region;

KW monoclonal antibody; MAb; binding affinity; EDTA; DOTA; tumour; cancer;

KW colorectal; breast; metal chelate; haptens; ss.

XX

OS Synthetic.

XX

PN AU9350602-A.

XX

PD 26-MAY-1994.

XX

PF 10-NOV-1993; 93AU-00050602.

XX

PR 12-NOV-1992; 92US-00975230.

XX

PA (HYBR-) HYBRITECH INC.

XX

PI Ahnweiler PM, Moore MD;

XX

DR WPI; 1994-209063/26.

DR

DR P-FSDB; AAR54136.

XX

PT Polypeptide used in imaging and treatment of carcinomas and tumours -

comprising substd antibody CDR having binding affinity for metal chelate of EDTA or DETA or analogues.

Disclosure; Fig 1B; 61pp; English.

The sequences given in AA068744-45 encode the light and heavy chain variable regions of the antibody designated CHA255. CHA255 is a murine monoclonal antibody (MAb) which is capable of binding metal complexes. Mutagenesis of the complementarity determining regions (CDRs) this MAb, causes the production of polypeptides with a particularly high binding affinity for EDTA or DOTA metal complexes. CDR1 and -3 of the heavy chain, and CDR2 and -3 of the light chain were targeted for mutagenesis. Five residues of both CDR1 and -3 of the CHA255heavy chain, five of seven residues of light chain CDR and six of nine light chain CDR3 residues were specifically targeted for codon-based mutagenesis. The mutagenised MAb's can be used in compositions for in vivo imaging of malignant tissues or tumours. They are also useful for the treatment of malignant tissues or tumours eg. colorectal or breast cancer. Both methods involve the use of radionuclides which bind to metal chelates or haptens which are specifically delivered to the target site by a targeting molecule. CDR derived peptides may be used to construct bi-functional antibodies having dual specificities, or as donor or recipients of CDR sequences

Sequence 330 BP; 88 A; 85 C; 80 G; 77 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 345 Length: 330
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-5 (1-6) x AA068744 (1-330)

Qy 1 GlyTyr***ValGluGlu 6

Db 293 GGGTATTTCGGTGGAGGAA 310

RESULT 5

ABX64662/C

ID ABX64662 standard; cDNA; 384 BP.

AC ABX64662;

DT 03-MAR-2003 (first entry)

XX Human gene trapped sequence (GTS) cDNA SEQ ID NO 269.

XX Human; gene trapped sequence; GTS; gene discovery; forensic; mapping; gene therapy; antisense therapy; prenatal analysis; autoimmune disease; developmental disorder; aging; cancer; Crohn's disease; schizophrenia; multiple sclerosis; skin disorder; rheumatoid arthritis; skeletal system; atherosclerosis; cardiovascular disease; degenerative disease; neural; Alzheimer's disease; osteoporosis; asthma; infection; ss.

OS Homo sapiens.

XX US2002110809-A1.

XX 15-AUG-2002.

XX 28-APR-2000; 2000US-00560863.

XX 30-APR-1999; 99US-0132408P.

XX (NEHL/) NEHLS M. C.

XX (ZAMB/) ZAMBROWICZ B.

XX (SAND/) SANDS A. T.

XX Nehls MC, Zambrowicz B, Sands AT;

XX WPI; 2003-090170/08.

XX Novel human polynucleotides that correspond to human gene trapped sequences useful for gene discovery, gene therapy, as markers for gene expression analysis, forensic analysis and determining genetic basis of diseases.

XX Claim 1; SEQ ID NO 269; 36pp; English.

XX This invention describes isolated or purified polynucleotides corresponding to human gene trapped sequences (GTS) comprising a sense or antisense sequence chosen from 1000 sequences or its fragment of 8 contiguous nucleotides. GTSs are useful for gene discovery and as markers for gene expression analysis, identifying and mapping the coding regions of the mammalian and particularly human, genome, forensic analysis, for cross species hybridisation analysis, genetic manipulation, antisense inhibition, gene targeting, identification or generation of full-length cDNA, mapping the human genome, gene or antisense therapy, gene delivery and determining the genetic basis of human disease. Portion of the GTS sequences are useful as a hybridisation probe or for chromosome mapping, and can be incorporated into phage display system that can be used to screen for proteins, or other ligands, that are capable of binding an amino acid sequence encoded by the GTS sequences. The GTS sequences are also useful to regulate gene expression, as a part of ribozyme and/or triple helix sequences that can be used to regulate gene expression, as components of diagnostic methods, for analysing single nucleotide polymorphisms, and also as genetic markers for prenatal analysis of congenital traits or defects. The polynucleotides of the invention are useful for diagnosis, prognosis of disorders involving developmental and differentiation processes and for the identification of subjects having a predisposition to such disorders. Diseases or natural processes that can be correlated with the expression of mutant or normal, variants of GTS include, aging, cancer, autoimmune disease, Crohn's disease, multiple sclerosis, immune disorders, schizophrenia, skin disorders, rheumatoid arthritis, atherosclerosis, cardiovascular disease, degenerative diseases of the neural or skeletal systems, Alzheimer's disease, osteoporosis, asthma, and infections. GTSs can also be used to identify the specific locations of exon splice junctions, which are important in the study of disease and cancer. Modulating the level of expression of one or more genes and/or regulating activity of one or more peptides or proteins is useful for modifying development and cell differentiation and treating development and cell differentiation disorders. ABX64402-ABX65401 represent the human gene trapped cDNA sequences described in the disclosure of the invention. NOTE: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20020110809

XX Sequence 384 BP; 96 A; 91 C; 87 G; 105 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.: 407 Length: 384
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x ABX64662 (1-384)

Qy 1 GlyTyr***ValGluGlu 6

Db 328 GGGTACAGTGTAGAAGAG 311

RESULT 6

ABV07265

ID ABV07265 standard; cDNA; 400 BP.

XX AC ABV07265;

XX 13-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 7256.

XX

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
OS Homo sapiens.
PN WO200160860-A2.
XX 23-AUG-2001.
PD 20-FEB-2001; 2001WO-US005171.
PF 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
DR Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 1177; 11750pp; English.
PS The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
SQ Sequence 400 BP; 91 A; 103 C; 85 G; 121 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 426 Length: 400
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0
US-10-030-194A-5 (1-6) x ABV07265 (1-400)
Qy 1 GlyTyr***ValGluGlu 6
Db 25 GGATATGCAGTAGAGGAA 42
RESULT 7
ACH17756
ID ACH17756 standard; cDNA; 417 BP.
XX ACH17756;
AC 13-OCT-2003 (first entry)
XX Human adult heart cDNA #2070.
DE Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
OS

XX
FN US2003073623-A1.
XX 17-APR-2003.
XX 30-JUL-2001; 2001US-00918995.
XX 30-JUL-2001; 2001US-00918995.
PR (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
DR New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX Claim 1; SEQ ID NO 4968; 44pp; English.
PS The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030073623
XX SQ Sequence 417 BP; 123 A; 120 C; 92 G; 82 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 445 Length: 417
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0
US-10-030-194A-5 (1-6) x ACH17756 (1-417)
Qy 1 GlyTyr***ValGluGlu 6
Db 314 GGATATGCTGTAGAGGAG 331
RESULT 8
ABV37205
ID ABV37205 standard; cDNA; 431 BP.
XX ABV37205;
AC 16-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 37196.
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX

OS Homo sapiens.
XX WO200160860-A2.
PN 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
PF 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 7649; 11750pp; English.
PS The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 431 BP; 100 A; 110 C; 94 G; 127 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 462 Length: 431
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x ABV37205 (1-431)

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 55 GGATATGCAGTAGAGGAA 72

RESULT 9
ID ABA13307
XX ABA13307 standard; cDNA; 460 BP.
AC ABA13307;
XX 23-JAN-2002 (first entry)
DT Human nervous system related polynucleotide SEQ ID NO 2314.
DE Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antischlicking; antianaemic; antiarthritis; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.

XX Homo sapiens.
OS WO200159063-A2.
PN 16-AUG-2001.
XX 17-JAN-2001; 2001WO-US001334.
PF 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI: 2001-541565/60.
DR P-PSDB; ABB16981.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX Claim 1; SEQ ID NO 2314; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 460 BP; 122 A; 59 C; 89 G; 188 T; 0 U; 2 Other;

Alignment Scores:
Pred No.: 496 Length: 460
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x ABA13307 (1-460)

Qy 1 GlyTyr***ValGluGlu 6
Db 27 GGCTACAGTGTGAGAA 44

RESULT 10
ABK77009
ID ABK77009 standard; DNA; 479 BP.
XX
AC ABK77009;
XX
DT 13-AUG-2002 (first entry)
XX
DE Bacillus licheniformis genomic sequence tag (GST) #4300.
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
XX physiological provocation; ds.
OS Bacillus licheniformis.
XX WO200229113-A2.
XX
XX 11-APR-2002.
XX
XX 05-OCT-2001; 2001WO-US031437.
XX
XX 06-OCT-2000; 2000US-00680598.
XX
XX 27-MAR-2001; 2001US-0279526P.
XX
XX (NOVO) NOVOZYMES BIOTECH INC.
XX

26-SEP-2000; 2000US-0235484P.
27-SEP-2000; 2000US-0235834P.
27-SEP-2000; 2000US-0235836P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239935P.
13-OCT-2000; 2000US-0239937P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241786P.
20-OCT-2000; 2000US-0241787P.
20-OCT-2000; 2000US-0241808P.
20-OCT-2000; 2000US-0241809P.
20-OCT-2000; 2000US-0241826P.
20-OCT-2000; 2000US-024221P.
01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249264P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250391P.
01-DEC-2000; 2000US-0251160P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
11-DEC-2000; 2000US-0251990P.
05-JAN-2001; 2000US-0254097P.
2001US-0259678P.

PA (NOVO) NOVOZYMES AS.
XX
PI Berka R, Clausen IG;
XX
XX WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX
XX
PS Claim 4; SEQ ID NO 4300; 200pp; English.
XX
XX The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC -up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 479 BP; 152 A; 90 C; 114 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 518 Length: 479
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x ABK77009 (1-479)
Qy 1 GlyTyr***ValGluGlu 6
Db 334 GGATACAGTGTCTGAAGAG 351

RESULT 11
ABI99741
ID ABI99741 standard; cDNA; 490 BP.
XX
XX
AC ABI99741;
XX
XX 07-MAR-2002 (first entry)
DT
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:807.
XX
XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasoapastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
XX Mus musculus.
OS
XX
XX WO200188188-A2.
PN
XX
XX 22-NOV-2001.
PD
XX
XX 18-MAY-2001; 2001WO-JP004192.
PF
XX
XX 18-MAY-2000; 2000JP-00145977.
PR

XX
PA (UWNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
XX WPI; 2002-034733/04.
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX
XX Claim 2; Page 1987; 2690pp; English.
XX
XX The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasoapastic ischaemia) by measuring the
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
SQ Sequence 490 BP; 112 A; 101 C; 110 G; 167 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 532 Length: 490
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x ABI99741 (1-490)
Qy 1 GlyTyr***ValGluGlu 6
Db 344 GGCTATGCTGTGGAAGAG 361

RESULT 12
ABK62521/c
ID ABK62521 standard; cDNA; 498 BP.
XX
XX ABK62521;
XX
XX 18-JUN-2002 (first entry)
DT
DE Rat sequence differentially expressed in response to a hepatotoxin #428.
XX
XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KW differential expression; centrilobular necrosis; steatosis.
XX
XX Rattus norvegicus.
OS
XX
XX WO200210453-A2.
PN
XX
XX 07-FEB-2002.
PD
XX
XX 30-JUL-2001; 2001WO-US023872.
PF
XX
XX 31-JUL-2000; 2000US-0222040P.
PR
XX
XX 02-NOV-2000; 2000US-0244880P.
PR
XX
XX 11-MAY-2001; 2001US-0290029P.
PR
XX
XX 15-MAY-2001; 2001US-0290645P.
PR
XX
XX 22-MAY-2001; 2001US-0292336P.
PR
XX
XX 06-JUN-2001; 2001US-0295798P.
PR
XX
XX 13-JUN-2001; 2001US-0297457P.
PR

PR 19-JUN-2001; 2001US-0298884P.
PR 09-JUL-2001; 2001US-0303459P.
XX (GENE-) GENE LOGIC INC.
PA Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX WPI; 2002-241625/29.
DR
XX
XX Predicting toxic effects of compounds or the progression of these toxic
PT effects by determining the changes in gene expression in tissues or cells
PT exposed to the toxin and comparing these to gene expression in unexposed
PT tissues or cells.
XX
PS Claim 1; SEQ ID NO 428; 239pp; English.
XX
CC The invention relates to methods for predicting toxic effects of
CC compounds or the progression of these toxic effects by determining the
CC global changes in gene expression in tissues or cells exposed to the
CC toxin and comparing these to gene expression in unexposed tissues or
CC cells. Also included are methods of predicting at least one toxic effect
CC of a compound or progression of a toxic effect, preferably the
CC hepatotoxicity of a compound, comprising detecting the level of
CC expression in a tissue or cell sample exposed to the compound of two or
CC more genes listed in the specification, where differential expression of
CC the genes is indicative of at least one toxic effect or progression. The
CC method can also be used to identify an agent which modulates the toxic
CC response and predict cellular pathways that a compound modulates in a
CC cell. The methods utilize a set of at least two probes (on a solid
CC support in kit form), where each of the probes comprises a sequence that
CC specifically hybridizes to a gene listed in the specification, a computer
CC system comprising a database containing information identifying the
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC set of genes comprising at least two genes listed in the specification,
CC and a user interface to view the information used to present information
CC identifying the expression level in a tissue or cell of at least one gene
CC listed in the specification. The method is useful for elucidating global
CC changes in gene expression and for identifying toxicity markers in
CC tissues or cell exposed to a known toxin. The genes may be used as
CC toxicity markers in drug screening and toxicity assays. The genes and
CC gene expression information may be used as diagnostic markers for the
CC prediction or identification of the physiological state of tissue or cell
CC sample that has been exposed to a compound or agent. Hepatotoxicity is
CC characterised by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC which is differentially expressed in response to a hepatotoxic agent
XX
SQ Sequence 498 BP; 135 A; 127 C; 119 G; 116 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 541 Length: 498
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x ABK62521 (1-498)

QY 1 GlyTyr***ValGluGlu 6
|||||
DB 245 GGCTACTCTGTGAAGAG 228

RESULT 13
ADB56048/c
ID ADB56048 standard; DNA; 498 BP.
XX
AC ADB56048;
XX
XX 04-DEC-2003 (first entry)
DT
XX Toxicity-related gene, SEQ ID 1074.
DE
XX

KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW drug screening; toxicity assay; db.
XX
XX Unidentified.
XX
XX WO2003064624-A2.
XX
XX 07-AUG-2003.
XX
XX 31-JAN-2003; 2003WO-US003194.
XX
XX 31-JAN-2002; 2002US-00060087.
PR 15-MAR-2002; 2002US-0364045P.
PR 15-MAR-2002; 2002US-0364045P.
PR 30-DEC-2002; 2002US-0436643P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higge B, Castle A, Elashoff M;
XX WPI; 2003-689530/65.
XX
PT Predicting a toxic effect of a compound, useful in identifying toxicity
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to
XX the compound.
XX
PS Claim 1; SEQ ID NO 1074; 1156pp; English.
XX
CC The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 498 BP; 135 A; 127 C; 119 G; 116 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 541 Length: 498
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 10 Gaps: 0

US-10-030-194A-5 (1-6) x ADB56048 (1-498)

QY 1 GlyTyr***ValGluGlu 6
|||||
DB 245 GGCTACTCTGTGAAGAG 228

RESULT 14
ADC75197/c
ID ADC75197 standard; cDNA; 533 BP.
XX
AC ADC75197;
XX
XX 01-JAN-2004 (first entry)
DT
XX T harzianum phytopathogen resistance-related contig cDNA - SEQ ID 121.
DE
XX

KW rice; Yeast; poppy; plant; disease resistance; anti-fungal;
XX phytopathogen; gene shuffling; ss.

OS Hypocrea lixii.

XX WO2003020905-A2.

XX 13-MAR-2003.

XX 30-AUG-2002; 2002WO-US027883.

XX 31-AUG-2001; 2001US-0316392P.

XX (DOWC) DOW CHEM CO.

XX Shukla V, Butler H, Larrinua I, Reddy AS;

XX WPI; 2003-290185/28.

XX Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
PT sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
PT rhoheas, useful for conferring disease resistance in plants.

XX Claim 1; SEQ ID NO 121; 617pp; English.

CC The invention relates to a novel isolated nucleic acid derived from
CC Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
CC (yeast), Trichoderma harzianum (Hypocrea lixi) and Papaver rhoheas
CC (poppy) and a sequence that hybridises to them under conditions of low
CC stringency, where expression of the nucleic acid in a plant results in a
CC disease resistance phenotype. The polynucleotides of the invention
CC demonstrate anti-fungal activity and may be useful in conferring disease
CC resistance in a plant against phytopathogen such as Aspergillus flavus,
CC Gibberella fujikuroi and Gibberella zeae. Furthermore, the
CC polynucleotides may be useful to retrieve unknown sequences and in gene
CC shuffling or sexual PCR procedures. The current sequence is that of the
CC phytopathogen resistance-related contig cDNA of the invention.

XX Sequence 533 BP; 136 A; 124 C; 116 G; 157 T; 0 U; 0 Other;

Alignment Scores: 583 Length: 533
Pred. No.: 27.00 Matches: 5
Score: 83.33% Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 96.43% Indels: 0
Query Match: 10 Gaps: 0
DB:

US-10-030-194A-5 (1-6) x ADC75197 (1-533)

Qy 1 GlyTyr***ValGluGlu 6

Db 126 GGATACACGTCGAGGAA 109

RESULT 15

ABK73924

ID ABK73924 standard; DNA; 535 BP.

XX ABK73924;

XX 13-AUG-2002 (first entry)

XX Bacillus licheniformis genomic sequence tag (GST) #1215.

XX Differential gene expression; genomic sequenced tag; GST;

XX altered culture condition; environmental stress;

XX physiological provocation; ds.

XX Bacillus licheniformis.

XX WO200229113-A2.

XX 11-APR-2002.

XX 05-OCT-2001; 2001WO-US031437.

XX 06-OCT-2000; 2000US-00680598.

XX 27-MAR-2001; 2001US-0279526P.

XX (NOVO) NOVOZYMES BIOTECH INC.

XX (NOVO) NOVOZYMES AS.

XX Berka R, Clausen IG;

XX WPI; 2002-416684/44.

XX Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.

XX Claim 4; SEQ ID NO 1215; 200pp; English.

CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions.
CC environmental stress or other physiological provocation. Extensive follow
CC up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 535 BP; 172 A; 108 C; 161 G; 93 T; 0 U; 1 Other;

Alignment Scores: 585 Length: 535
Pred. No.: 27.00 Matches: 5
Score: 83.33% Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 96.43% Indels: 0
Query Match: 6 Gaps: 0
DB:

US-10-030-194A-5 (1-6) x ABK73924 (1-535)

Qy 1 GlyTyr***ValGluGlu 6

Db 373 GGATATTCAGTGGAGAA 390

RESULT 16

ABV59077

ID ABV59077 standard; cDNA; 565 BP.

XX ABV59077;

XX 13-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 59068.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

PN WO2000160860-A2.
XX
PD 23-AUG-2001.
XX
PF
XX
PR 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
XX Schlegel R, Endege WO, Monahan JE;
PI
XX
XX WPI; 2001-662795/76.
DR
XX
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 11327; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 565 BP; 145 A; 131 C; 122 G; 167 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 621 Length: 565
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0
US-10-030-194A-5 (1-6) x ABV59077 (1-565)
Qy 1 GlyTyr***ValGluGlu 6
Db 209 GGATATGCAGTAGAGAA 226
RESULT 17
ACF68403
ID ACF68403 standard; DNA; 696 BP.
XX
XX ACF68403;
AC
XX
XX 20-NOV-2003 (first entry)
DT
XX
XX Photorhabdus luminescens nucleotide sequence #6870.
DE
XX
XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough; gene; ds.
XX
XX Photorhabdus luminescens.
OS
XX
XX WO200294867-A2.
PN
XX

PD 28-NOV-2002.
XX
XX 07-FEB-2002; 2002WO-IB003040.
PF
XX
XX 07-FEB-2001; 2001FR-00001659.
PR
XX
XX (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
PI
XX
XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI Buchrieser C;
XX
XX WPI; 2003-148459/14.
DR
XX
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
PT
XX
XX Claim 2; SEQ ID NO 6870; 1205pp; French.
PS
XX
XX The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX
XX Sequence 696 BP; 226 A; 116 C; 131 G; 223 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 781 Length: 696
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 10 Gaps: 0
US-10-030-194A-5 (1-6) x ACF68403 (1-696)
Qy 1 GlyTyr***ValGluGlu 6
Db 169 GGGTATACCGTTGAAGAA 186
RESULT 18
AAZ15431/C
ID AAZ15431 standard; cDNA; 708 BP.
XX
XX AAZ15431;
AC
XX
XX 12-OCT-1999 (first entry)
DT
XX
XX Human gene expression product cDNA sequence SEQ ID NO:2900.
DE
XX
XX Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
XX Homo sapiens.
OS
XX

PN W09938972-A2.
XX
PD 05-AUG-1999.
XX
XX 28-JAN-1999; 99WO-US001619.
PF 28-JAN-1998; 98US-0072910P.
PR 24-FEB-1998; 98US-0075954P.
PR 31-MAR-1998; 98US-0080114P.
PR 03-APR-1998; 98US-0080515P.
PR 03-APR-1998; 98US-0080666P.
PR 21-OCT-1998; 98US-0105234P.
PR 28-OCT-1998; 98US-0105877P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lanson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX
DR WPI; 1999-494092/41.
XX
XX Novel human genes and their expression products which are differentially
PT expressed in different cell types.
PT
XX
XX Claim 1; Page 1400-1401; 2479pp; English.
XX
XX The present invention describes a library of human polynucleotides
CC comprising the sequences given in AA212532 to AA217779. Also described is
CC a method of detecting differentially expressed genes correlated with the
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one of
CC the 5248 polynucleotide sequences given in AA212532 to AA217779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists
XX
SQ Sequence 708 BP; 166 A; 168 C; 135 G; 233 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 796 Length: 708
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-5 (1-6) x AA215431 (1-708)

Qy 1 GlyTyr***ValGluGlu 6
Db 413 GGTATAGTGTGAAGAG 396

RESULT 19
AAK90974/C
ID AAK90974 standard; DNA; 747 BP.
XX
AC AAK90974;
XX

DT 05-NOV-2001 (first entry)
XX Human digestive system antigen genomic sequence SEQ ID NO: 4550.
DE
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
OS Homo sapiens.
XX
XX WO200155314-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001324.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 30-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220984P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225113P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-023497P.
PR 25-SEP-2000; 2000US-023497P.
PR 26-SEP-2000; 2000US-023498P.
PR 26-SEP-2000; 2000US-023484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 11-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
XX Disclosure; SEQ ID NO 4550; 986pp; English.
XX The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention
XX
SQ Sequence 747 BP; 242 A; 160 C; 150 G; 195 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 844 Length: 747
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 4 Gaps: 0
US-10-030-194A-5 (1-6) x AAK90974 (1-747)
QY 1 GlyTyr***ValGluGlu 6
|||
Db 203 GGTATTCTGTGGAAGAA 186
RESULT 20
AAS32009/c
ID AAS32009 standard; DNA; 747 BP.
XX
AC AAS32009;
XX
DT 04-DEC-2001 (first entry)
XX Human liver associated genomic DNA #183.
XX
XX Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;
KW Chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW anirhematic; antiproliferative; cytoskeletal; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
XX anti-infertility.
XX Homo sapiens.
OS
XX WO200155355-A1.
PN
XX
XX
PD
XX
XX
PF 17-JAN-2001; 2001WO-US001351.
XX

PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249285P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-457728/49.

Isolated nucleic acid molecule encoding a human liver related protein is used in preventing, treating or ameliorating disorders of the liver particularly cancer of the liver.

XX Claim 1; SEQ ID NO 485; 526pp; English.

PS Sequences AAG31827-AAS32182 represent genomic DNA molecules, which encode

XX the liver associated polypeptides of the invention. Liver associated

CC polypeptides and their associated polynucleotides are useful in the

CC diagnosis, treatment and prevention of various types of disorders in e.g.

CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A

CC pathological condition can be determined by detecting the presence or

CC absence of a mutation in a liver associated polynucleotide. The treatable

CC disorders include autoimmune diseases such as rheumatoid arthritis,

CC hyperproliferative disorders such as neoplasms of the breast or liver,

CC cardiovascular disorders such as cardiac arrest, cerebrovascular

CC disorders such as cerebral ischaemia, nervous system disorders such as

CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,

CC ocular disorders such as corneal infection, endocrine disorders such as

CC premature labour and infertility, gastrointestinal disorders such as

CC Crohn's disease, renal disorders such as glomerulonephritis and

CC respiratory disorders such as asthma and pleurisy. The polypeptides can

CC also be used to aid wound healing, to prevent skin aging due to sunburn,

CC to maintain organs before transplantation, to regenerate tissues and in

CC chemotaxis. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 747 BP; 242 A; 160 C; 150 G; 195 T; 0 U; 0 Other;

Alignment Scores:

Pred No.:	844	Length:	747
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	5	Gaps:	0

US-10-030-194A-5 (1-6) x AAS32009 (1-747)

Qy 1 GlyTyr***ValGluGlu 6

Db 203 GGTATTCTGTGGAAGAA 186

RESULT 21

ID ABN90364/c

ABN90364 standard; DNA; 747 BP.

XX

AC ABN90364;

XX

DT 24-JUL-2002 (first entry)

XX

DE Human liver antigen HLPBW64 genomic sequence, SEQ ID NO:485.

XX

XX Human; liver antigen; liver disorder; hepatic disorder; infection;

KW hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;

KW cirrhosis; granulomatous hepatitis; toxin damage; drug damage;

KW autoimmune disease; Wilson's disease; primary biliary cirrhosis;

KW neoplastic disorder; cancer; tumour; portal hypertension;

KW gastrointestinal disorder; hepatitis; drug screening; gene therapy;

KW chromosome mapping; forensic analysis; antibody preparation;

KW hepatotropic; cytostatic; antiinflammatory; virucide; antibacterial;

KW fungicide; parasiticide; antidote; immunosuppressive; gene; ds.

XX

OS Homo sapiens.

XX

OS US2002042096-A1.

PN

XX

FD 11-APR-2002.

XX

XX 17-JAN-2001; 2001US-00764887.

PF

XX

XX 31-JAN-2000; 2000US-0179065P.

PR

XX 04-FEB-2000; 2000US-0180628P.

PR

XX 28-JUN-2000; 2000US-0214886P.

PR

XX 07-JUL-2000; 2000US-0216647P.

PR

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 22-AUG-2000; 2000US-0225758P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 08-SEP-2000; 2000US-0231413P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 27-SEP-2000; 2000US-0235834P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236388P.

PR 29-SEP-2000; 2000US-0236399P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241809P.

PR 01-NOV-2000; 2000US-0244617P.

PR 17-NOV-2000; 2000US-0249299P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.

XX

PA (ROSE/) ROSEN C A.

PA (RUBE/) RUBEN S M.

PA (BARA/) BARASH S C.

XX

PI Rosen CA, Ruben SM, Barash SC;

XX

XX WPI; 2002-381944/41.

XX

PT New nucleic acid encoding human liver antigens, useful for diagnosis,

PT treatment and prevention of e.g. hepatitis and hepatic cancer, also

XX related polypeptides and antibodies.

PS Disclosure; SEQ ID NO 485; 181pp; English.

XX

CC The invention relates to 145 novel human liver antigens (ABP40831-

CC ABP40975) and to cDNAs encoding them (ABN90036-ABN90180), and also

CC encompasses polypeptides 90% identical and polynucleotides 95% identical

CC to the sequences of the invention. The invention additionally relates to

CC recombinant vectors and host cells comprising human liver antigen

CC polynucleotides, antibodies against human liver antigens, and the use of

CC liver antigen polynucleotides and polypeptides in diagnosing, treating,

CC prognosing or preventing various disorders of the liver. Such conditions

CC include viral infections (e.g., cytomegalovirus, Epstein-Barr virus,

CC hepatitis A virus, hepatitis B virus and hepatitis C virus), parasitic

CC infections (e.g., Clonorchis sinensis, Echinococcus granulosus and

CC Entamoeba histolytica), and also bacterial and fungal infections. Other

disorders that may be treated include inflammatory conditions (e.g., cirrhosis and granulomatous hepatitis), damage caused by drugs or toxins, autoimmune diseases (e.g. Wilson's disease, Primary biliary cirrhosis), neoplastic disorders (e.g., adenomas, haemangiomas and hepatocellular carcinoma), portal hypertension, or gastrointestinal disorders (e.g., peptic ulcers, gastritis and peritoneal diseases). Liver antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate liver antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as molecular weight markers or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. Sequences ABN90182-ABN90537 represent human liver antigen genomic sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO at seqdata.uspto.gov/sequence/

SQ Sequence 747 BP; 242 A; 160 C; 150 G; 195 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 844 Length: 747
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x ABN90364 (1-747)

Qy 1 GlyTyr***ValGluGlu 6
 |||||
 Db 203 GGTATTCTGTGGAAGAA 186

RESULT 22

ADJ15277/c

ID ADJ15277 standard; DNA; 747 BP.

XX AC ADJ15277;

XX DT 20-MAY-2004 (first entry)

XX DE Human liver-related genomic DNA - SEQ ID 485.

KW liver; virucide; fungicide; antibacterial; antiparasitic; hepatotropic;
 KW antiinflammatory; cytostatic; litholytic; antirheumatic; antiarthritic;
 KW neuroprotective; antidiabetic; anticoagulant; thrombolytic;
 KW antiarteriosclerotic; cardiant; haemostatic; antiarrhythmic;
 KW ophthalmological; antiarteriosclerotic; vasotropic; osteopathic;
 KW neotropic; antiparkinsonian; anticonvulsant; neuroleptic; vasotropic;
 KW cytostatic; gynaecological; viral; fungal; bacterial;
 KW parasitic infection; cirrhosis; Wilson's disease;
 KW gastrointestinal disorder; pancreatic; gallbladder; immune; blood;
 KW hyperproliferative; cardiovascular; respiratory; musculoskeletal system;
 KW neurological; endocrine; reproductive system; developmental; inherited;
 human; ds.

XX OS Homo sapiens.

XX FN US2003077602-A1.

XX PD 24-APR-2003.

XX PF 14-FEB-2002; 2002US-00073961.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 16-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 18-AUG-2000; 2000US-0225759P.
 PR 22-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 30-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.

the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 834 BP; 218 A; 221 C; 221 G; 174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 952 Length: 834
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x ACA18963 (1-834)

Qy 1 GlyTyr***ValGluGlu 6
Db 336 GGTACGCTGTCGAGGAA 319

RESULT 24
ADC91913/c
ID ADC91913 standard; DNA; 849 BP.

XX ADC91913;
XX
DT 01-JAN-2004 (first entry)
XX
DE E. faecium DNA sequence SEQ ID 1540.

XX ds; gene; urinary tract infection; bacteraemia; endocarditis; wound;
KW abdominal-pelvic infection.

XX Enterococcus faecium.

XX US6583275-B1.

XX 24-JUN-2003.

XX 30-JUN-1998; 98US-00107532.

XX 02-JUL-1997; 97US-0051571P.

XX 14-MAY-1998; 98US-0085598P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Doucette-Stamm LA, Bush D;
XX
DR WPI: 2003-799836/75.
DR P-PSDB; ADC95567.
XX
PT New isolated nucleic acid derived from *Enterococcus faecium* encoding an
PT *Enterococcus faecium* polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
PT infection.

XX Example 1; SEQ ID NO 1540; 243pp; English.

XX The invention relates to an isolated nucleic acid derived from
XX *Enterococcus faecium* encoding an *Enterococcus faecium* polypeptide having
XX one of 10 fully defined sequences given in the (or comprising 40
XX sequential nucleotides chosen from any of the nucleic acids, its
XX complement or sequences hybridising to it). Also included are a
XX recombinant vector comprising the nucleic acid operably linked to a
XX transcription regulatory element, a cell comprising the vector and a
XX single-stranded probe comprising the nucleic acid. The nucleic acids are
XX chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
XX The nucleic acids are useful for diagnosing pathological conditions
XX resulting from *E. faecium* bacterial infection (e.g. urinary tract
XX infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
XX infection) and for screening drugs such as agonists and antagonists. The
XX nucleic acid is useful for recombinant production of *Candida albicans* -
XX derived peptides or antisense polypeptides. Pharmaceutical compositions
XX and vaccines containing the nucleic acid are useful for preventing or
XX treating *Enterococcus faecium* infections. The present sequence represents
XX one if the disclosed *E. faecium* nucleic acids.

XX Sequence 849 BP; 310 A; 120 C; 157 G; 262 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 971 Length: 849
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 10 Gaps: 0

US-10-030-194A-5 (1-6) x ADC91913 (1-849)

Qy 1 GlyTyr***ValGluGlu 6
Db 694 GGATATAGTGTGGAGAG 677

RESULT 25
AAH31919/c
ID AAH31919 standard; DNA; 963 BP.

XX AAH31919;

XX 30-JUL-2001 (first entry)

XX Human olfactory receptor polynucleotide, SEQ ID NO: 492.

XX Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.

XX Homo sapiens.

XX WO200127158-A2.

XX 19-APR-2001.

XX 06-OCT-2000; 2000WO-US027582.

XX 08-OCT-1999; 99US-0158615P.

XX 24-FEB-2000; 2000US-0184809P.

XX (DIG1-) DIGISCENTS.

XX

(SENO-) SENOMYX INC.

PR 02-JUN-2000; 2000US-0208834P.
 PR 02-JUN-2000; 2000US-0208861P.
 PR 07-JUN-2000; 2000US-0209868P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Patterson C, Tribouley CM, Yao MG, Griffin JA, Thornton M, Lu Y;
 PI Kallick DA, Gandhi AR, Au-Young J;
 XX
 DR WPI; 2002-106199/14.
 DR P-PSDB; AAU80513.
 XX
 XX New G-protein coupled receptors useful for treating or preventing cell
 PT proliferative (e.g. leukemia), neurological (e.g. stroke), cardiovascular
 PT or autoimmune/inflammatory disorders.
 XX
 XX Claim 5; Page 147; 148pp; English.
 XX
 CC The invention relates to a novel human G-coupled receptor (I). (I) and
 CC its corresponding polynucleotides are useful for diagnosing, treating or
 CC preventing cell proliferative diseases (e.g. lymphoma, leukaemia, breast
 CC cancer or cirrhosis), neurological disorders (e.g. stroke, Alzheimer's
 CC disease, multiple sclerosis or mental retardation), cardiovascular
 CC diseases (e.g. atherosclerosis, angina pectoris or congestive heart
 CC failure), gastrointestinal disorders (e.g. dysphagia, indigestion or
 CC gastritis), autoimmune/inflammatory disorders (e.g. AIDS, Crohn's disease
 CC or systemic lupus erythematosus) or metabolic disorders (e.g. diabetes or
 CC obesity), or viral infections (e.g. infection by herpesvirus or
 CC parvovirus). ABK16615-ABK16637 represent novel human G-coupled receptor
 CC coding sequences of the invention
 XX
 SQ Sequence 987 BP; 237 A; 253 C; 168 G; 329 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.14e+03 Length: 987
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 6 Gaps: 0
 US-10-030-194A-5 (1-6) x ABK16635 (1-987)
 Qy 1 GlyTyr***ValGluGlu 6
 Db 792 GGGTACACAGTAGAGGAA 775
 RESULT 28
 ABZ43050/c
 ID ABZ43050 standard; DNA; 987 BP.
 AC ABZ43050;
 XX
 XX 06-MAR-2003 (first entry)
 DT
 DE Human GPCR polynucleotide SEQ ID NO 361.
 XX
 XX Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
 KW drug development; gustatory; taste; fragrance; gene; ds.
 KW
 OS Homo sapiens.
 XX
 XX WO200216548-A2.
 PN
 XX 28-FEB-2002.
 PD
 XX 30-JUL-2001; 2001WO-IB001446.
 PF
 XX 04-AUG-2000; 2000JP-00237818.
 PR 13-FEB-2001; 2001JP-00034434.
 XX
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 PA
 XX

PI Haga T, Takeda S, Mitaku S;
 XX
 DR WPI; 2002-304118/34.
 DR P-PSDB; ABP95776.
 XX
 PT Database global search for G protein-coupled receptors, proteins and
 PT encoded genes for studying in vivo signal transduction mechanism and
 PT identifying targets for drug development.
 XX
 XX Claim 9; SEQ ID NO 361; 97pp + Sequence Listing; Japanese.
 XX
 CC The invention relates to a method for screening G protein-coupled
 CC receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-
 CC ABP95942) by extracting open-reading frames containing 6-8 transmembrane
 CC domains with 250-1000 amino acid residues to give a gene homologous with
 CC a known GPCR gene. The receptor proteins and encoded genes are useful for
 CC studying in vivo signal transduction mechanism and identifying targets
 CC for drug development e.g. based on olfactory and gustatory receptors in
 CC form of agonists and antagonists by screening intrinsic and extrinsic
 CC ligands as bitter taste inhibitors, taste enhancers and fragrance
 CC improvers. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 987 BP; 237 A; 253 C; 168 G; 329 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.14e+03 Length: 987
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 6 Gaps: 0
 US-10-030-194A-5 (1-6) x ABZ43050 (1-987)
 Qy 1 GlyTyr***ValGluGlu 6
 Db 792 GGGTACACAGTAGAGGAA 775
 RESULT 29
 ABK68526/c
 ID ABK68526 standard; DNA; 987 BP.
 AC ABK68526;
 XX
 XX 02-JUL-2002 (first entry)
 DT
 DE Human DNA for olfactory and pheromone G protein-coupled receptor #126.
 XX
 XX Human; ds; gene; olfactory and pheromone G protein coupled receptor;
 KW GPCR; tranquilizer; antidepressant; neuroleptic; endocrine; anabolic;
 KW anorectic; taste; fragrance; food additive; cosmetic; cell migration;
 KW sterility; psychotic disorder; neurological disorder; anxiety;
 KW schizophrenia; manic depression; depression; axonal growth;
 KW menstrual cycle; appetite sexual motivation; sexual attraction;
 KW aggression.
 KW
 OS Homo sapiens.
 XX
 XX WO200224726-A2.
 PN
 XX 28-MAR-2002.
 PD
 XX 21-SEP-2001; 2001WO-BE000162.
 PF
 XX 22-SEP-2000; 2000EP-00870211.
 PR
 XX (CHEM-) CHEMCOM SA.
 PA
 XX Veithen A;
 PI
 XX WPI; 2002-330013/36.
 DR

DR P-PSDB; AAU95639.

XX Novel pheromone G-protein coupled receptor and receptor-derived agonists, antagonists or inhibitors useful in food or cosmetic products or in the treatment or prevention of neurological disorders such as anxiety and schizophrenia.

PT

PT

XX Disclosure; Page 389-391; 833pp; English.

XX The invention relates to olfactory and Pheromone G-protein coupled receptor (GPCR) or a protein 95% identical to the GPCR, a specific active portion and its encoding polynucleotide. Also included are an agonist, antagonist or inhibitor of the GPCR or the polynucleotide, a vector comprising the polynucleotide, a cell transformed by the vector, a non-human mammal comprising a partial or total deletion of the polynucleotide encoding the receptor and screening (detection and possibly, recovering) of compounds which are known or not known to be agonist, antagonists or inhibitors of natural compounds to the GPCR. The receptor-derived agonists, antagonists, inhibitors or compounds are used as an improvement, elimination or substitution of an existing taste and/or a fragrance of (or in) the food and/or cosmetic products. They can also be used in the preparation of medicament in the treatment and/or prevention of a mammalian disorder, such as cell migration, sterility, psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, depression, for promoting axonal growth, nerve cell connection and nerve regeneration for modulating male and female endocrine functions, hormone production and the menstrual cycle, for the prevention or the treatment by stimulation of several mammalian behaviours, such as stimulation or suppression of appetite, sexual motivation, sexual attraction, aggression and for promoting or suppressing chemical communication between organisms. The present sequence is a human DNA encoding an olfactory and pheromone GPCR

XX Sequence 987 BP; 237 A; 253 C; 168 G; 329 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.14e+03	Length:	987
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	6	Gaps:	0

US-10-030-194A-5 (1-6) x ABK68526 (1-987)

OY 1 GlyTyr***ValGluGlu 6

DB 792 GGGTACACAGTAGAGGAA 775

RESULT 30

ABK37645/c

ID ABK37645 standard; CDNA; 987 BP.

XX

AC ABK37645;

XX

DT 08-MAY-2002 (first entry)

DE

DE DNA encoding G-coupled olfactory receptor #147.

XX Human; olfactory G-coupled receptor; sensory perception of odourant; odour composition; taste composition; gene; ss.

KW

OS Homo sapiens.

XX

PN WO200198526-A2.

XX

PD 27-DEC-2001.

XX

PF 22-JUN-2001; 2001WO-US020122.

XX

PR 22-JUN-2000; 2000US-0213812P.

XX

PR 13-MAR-2001; 2001US-00804291.

XX

PA (SENO-) SENOMYX INC.

XX

PI Zozulya S, Stryer L;

XX

DR WPI; 2002-083330/11.

DR P-PSDB; AAU85286.

XX

PT Representing sensory perception of one or more odorants for the identification and design of tastes and odors comprises providing a representative group of n olfactory receptors.

PT

XX Example; Page 124; 182pp; English.

XX The invention relates to a method of representing sensory perception of one or more odorants. The method comprises: (a) providing a representative class of n olfactory receptors or ligand binding domains (LBD) of these receptors; (b) measuring values X1 to Xn representative of at least one activity of one or more odorants selected from: (i) binding one or more odorants to the LBD of at least one of the n olfactory receptors; (ii) activating at least one of the n olfactory receptors with the one or more odorants; and (iii) blocking at least one of the n olfactory receptors with the one or more odorants; and (c) generating a representation of the sensory perception of odorants is useful for the design and formulation of odour and taste compositions. ABK37499-ABK37754 and ABK37918-ABK37921 represent human olfactory G-coupled receptor coding sequences and related PCR primers of the invention

XX Sequence 987 BP; 237 A; 253 C; 168 G; 329 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.14e+03	Length:	987
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	6	Gaps:	0

US-10-030-194A-5 (1-6) x ABK37645 (1-987)

OY 1 GlyTyr***ValGluGlu 6

DB 792 GGGTACACAGTAGAGGAA 775

RESULT 31

ADG83395/c

ID ADG83395 standard; DNA; 987 BP.

XX

AC ADG83395;

XX

DT 11-MAR-2004 (first entry)

DE

DE Human Olfactory and pheromone GPCR DNA #55.

XX

KW Olfactory; pheromone; GPCR; G protein-coupled receptor; ds; human; flavour; detergent; soap; shampoo; fragrance; appetite control; odour trap; receptor-mediated disorder; cell migration; cell death; cell growth; psychotic disorder; neurological disorder; anxiety; schizophrenia; maniac depression; depression; mound modification.

XX

OS Homo sapiens.

XX

PN US2003221205-A1.

XX

PD 27-NOV-2003.

XX

PF 13-MAR-2003; 2003US-00387629.

XX

PR 21-SEP-2001; 2001WO-BE000162.

XX

PA (VEIT/) VEITHEN A.

XX

PI Veithen A;

XX DR WPI; 2004-010900/01.

XX PT New isolated pheromone G-protein coupled receptor, or its active portion, useful for treating disorders affecting cell migration, death and growth, or psychotic and neurological disorders.

XX PT Disclosure; SEQ ID NO 109; 7pp; English.

XX PS

XX CC The invention relates to an isolated olfactory and pheromone G protein-coupled receptor (GPCR), or its active portion. Also included are an isolated polynucleotide encoding the GPCR, a vector comprising the polynucleotide, a cell comprising the vector, a non-human mammal comprising a deletion of the polynucleotide, a transgenic non-human mammal that overexpresses the polynucleotide, screening for compounds that bind to the receptor polypeptide molecules, identifying a compound as a ligand, agonist or antagonist of GPCR and a pharmaceutical composition comprising the receptor, vector or cell, and a carrier. The protein, polynucleotide, vector, cell and composition are useful for preventing or treating receptor-mediated disorders, e.g. disorders affecting cell migration, cell death, cell growth, or psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, depression or mood modification. They may also be useful in developing detergents, soaps, shampoo, fragrances, in appetite control and as an odour trap. The present sequence is an olfactory and pheromone G protein-coupled receptor DNA. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docID=20030221205.

XX CC

XX SQ Sequence 987 BP; 237 A; 253 C; 168 G; 329 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.:	1.14e+03	Length:	987
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	12	Gaps:	0

US-10-030-194A-5 (1-6) x ADG83395 (1-987)

Qy 1 GlyTyr***ValGluGlu 6

Db 792 GGGTACACAGTAGAGAA 775

RESULT 32

AD035088/c

ID AD035088 standard; DNA; 1001 BP.

XX AC AD035088;

XX DT 12-AUG-2004 (first entry)

XX DE Human KChIP1 SNP #49.

XX ds; diabetes; polymorphism;

KW voltage-gated potassium channel-interacting protein 1; KChIP1; diagnosis;

KW chromosome 5.

XX OS Homo sapiens.

XX PN WO2004041193-A2.

XX PD 21-MAY-2004.

XX PF 31-OCT-2003; 2003WO-US034681.

XX PR 01-NOV-2002; 2002US-0423545P.

XX PR 25-FEB-2003; 2003US-0449945P.

XX PR 09-JUN-2003; 2003US-0477111P.

XX PA (DECO-) DECODE GENETICS EHF.

XX PI Reynisdottir I, Gulcher JR, Grant SF, Thorleifsson G;

XX WPI; 2004-390509/36.

XX PT Diagnosing susceptibility of individual to type II diabetes, involves detecting polymorphism in voltage-gated potassium channel-interacting protein 1 nucleic acid.

XX PT

XX PS Claim 4; SEQ ID NO 162; 370pp; English.

XX CC The invention relates to a method of diagnosing (M1) a susceptibility to type II diabetes in an individual, by: (a) detecting a polymorphism in a voltage-gated potassium (Kv) channel-interacting protein 1 (KChIP1) nucleic acid; (b) detecting an alteration in the expression or composition of a polypeptide encoded by KChIP1 nucleic acid in a test sample; (c) determining the presence of a haplotype (H) comprising a haplotype designated A1-A6 and B1-B5 and haplotype of Icelandic population designated D2-D6 and Danish population designated D2-D5 as given in the specification at the 5q35 loci; (d) obtaining a nucleic acid sample, and analyzing for the presence of (H) at the 5q35 loci comprising a KChIP1 gene; (e) determining the presence of haplotype (H) comprising one or more markers and/or single nucleotide polymorphisms (SNPs) of 142 markers and/or SNPs (PM) as given in the specification in the locus on chromosome 5q35; or (f) screening for an at-risk haplotype in the KChIP1 nucleic acid that is more frequently present in an individual susceptible to type II diabetes compared to an individual who is not susceptible to type II diabetes. (M1) is useful for diagnosing susceptibility to type II diabetes in an individual. A host cell is useful for producing a polypeptide encoded by an isolated nucleic acid molecule having a polymorphism. An agent is useful for altering expression of a KChIP1 nucleic acid. The type II diabetes therapeutic agent is useful for the treatment of a disease or condition associated with KChIP1 in an individual. This sequence corresponds to a voltage-gated potassium (Kv) channel-interacting protein 1 (KChIP1) gene single nucleotide polymorphism (SNP) sequence.

XX SQ Sequence 1001 BP; 282 A; 240 C; 232 G; 246 T; 0 U; 1 Other;

Alignment Scores:

Pred. NO.:	1.16e+03	Length:	1001
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	12	Gaps:	0

US-10-030-194A-5 (1-6) x ADO35088 (1-1001)

Qy 1 GlyTyr***ValGluGlu 6

Db 93 GGATACACTGTTGAAGAA 76

RESULT 33

ABK97222/c

ID ABK97222 standard; cDNA; 1002 BP.

XX AC ABK97222;

XX DT 07-OCT-2002 (first entry)

XX DE Human G-protein coupled receptor (GPCR) GPCR gene #16.

XX G-protein coupled receptor; receptor; GPCR; GPCR; cardiomyopathy;

KW atherosclerosis; diabetes; cell signal processing; cancer; trauma;

KW metabolic pathway modulation; neuro-olfactory system; surgery;

KW neoplastic disorder; adenocarcinoma; lymphoma; prostate cancer;

KW uterus cancer; immune response; acquired immunodeficiency syndrome; AIDS;

KW asthma; Crohn's disease; multiple sclerosis;

KW Albright hereditary osteodystrophy; gene therapy; gene; ss.

XX OS Homo sapiens.

XX

PN WO200250117-A2.

XX 27-JUN-2002.

XX 18-DEC-2001; 2001WO-US049077.

XX 18-DEC-2000; 2000US-0256635P.

XX 21-DEC-2000; 2000US-0257876P.

XX 04-JAN-2001; 2001US-0259743P.

XX 10-JAN-2001; 2001US-0260718P.

XX 12-JAN-2001; 2001US-0261498P.

XX 24-JAN-2001; 2001US-0263689P.

XX 08-FEB-2001; 2001US-0267464P.

XX 22-FEB-2001; 2001US-0271021P.

XX 14-MAR-2001; 2001US-0275946P.

XX 23-MAR-2001; 2001US-0278150P.

XX 18-APR-2001; 2001US-0284591P.

XX 23-APR-2001; 2001US-0285718P.

XX 19-JUN-2001; 2001US-0299327P.

XX 16-AUG-2001; 2001US-0312902P.

XX (CURA-) CURAGEN CORP.

XX Padigar M, Kekuda R, Colman SD, Spytek KA, Ballinger RA;
PI Vernet CAM, Li L, Shenoy S, Casman SJ;XX WPI; 2002-528447/56.
XX P-PSDB; ABG68149.XX New G-protein coupled receptor polypeptides for treating or preventing
PT cardiomyopathy, atherosclerosis, diabetes, multiple sclerosis, acquired
PT immunodeficiency syndrome or cancer in humans.

XX Claim 5; Page 94; 110pp; English.

XX The present invention relates to a new G-protein coupled receptor (GPCRX)
CC polypeptide. The GPCRX polypeptide, GPCRX nucleic acid and antibody are
CC useful for treating, preventing or alleviating a GPCRX-associated
CC disorder or a pathological state in a subject, particularly a human. In
CC particular, the disorder is cardiomyopathy, atherosclerosis, diabetes, or
CC a disorder related to cell signal processing and metabolic pathway
CC modulation. The GPCRX polypeptide and nucleic acid are also useful for
CC diagnosing the presence of or predisposition to a disease associated with
CC altered levels of GPCRX, particularly cancer. The GPCRX nucleic acid and
CC polypeptides are especially useful in therapeutic or prophylactic
CC applications for disorders of the neuro-olfactory system, e.g. those
CC induced by trauma, surgery and/or neoplastic disorders. The DNA encoding
CC the protein is useful in gene therapy for treating the above conditions.
CC Furthermore, the nucleic acids and polypeptides are useful in treating
CC adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune
CC response, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's
CC disease, multiple sclerosis or Albright hereditary osteodystrophy. These
CC are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications. The present nucleic acid sequence represents one of a
CC collection (ABK97207-ABK97226) of human GPCRX genes that encode the human
CC GPCRX proteins (AAU68134-AAU68153) of the invention

XX SQ Sequence 1002 BP; 242 A; 256 C; 169 G; 335 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.16e+03	Length:	1002
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	6	Gaps:	0

US-10-030-194A-5 (1-6) x ABK97222 (1-1002)

Qy 1 GlyTyr***ValGluGlu 6

Db 798 GGGTACACAGTAGAGGAA 781

RESULT 34	
AAD60343/c	
ID	AAD60343 standard; DNA; 1002 BP.
XX	
XX	AAD60343;
XX	
DT	18-DEC-2003 (first entry)
XX	
DE	Human G-protein coupled receptor (GPCR) GPCR DNA #16.
XX	
KW	Human; G-protein coupled receptor; GPCR; stroke; obesity; virulence;
KW	multiple sclerosis; Alzheimer's disease; graft-versus-host disease;
KW	endometriosis; tissue typing; gene therapy; vaccine; antibacterial;
KW	retinal disease; infectious disease; bulimia; receptor; fungicide;
KW	Parkinson's disease; hypertension; acute heart failure; infection;
KW	cardiomyopathy; atherosclerosis; diabetes; detectability disorder;
KW	cancer; signal transduction pathway disorder; metabolic disorder;
KW	developmental disorder; osteoporosis; protozoacide; asthma; gene; de.
XX	
OS	Homo sapiens.
XX	
EH	Key Location/Qualifiers
FT	CDS 7..993
FT	/*tag= a
FT	/product= "Human GPCR protein"
FT	/note= "This CDS contains translation
FT	exceptions"
XX	
PN	US2003100491-A1.
XX	
PD	29-MAY-2003.
XX	
PF	18-DEC-2001; 2001US-00024399.
XX	
PR	18-DEC-2000; 2000US-0256635P.
PR	21-DEC-2000; 2000US-0257876P.
PR	04-JAN-2001; 2001US-0259743P.
PR	10-JAN-2001; 2001US-0260718P.
PR	12-JAN-2001; 2001US-0261498P.
PR	24-JAN-2001; 2001US-0263689P.
PR	08-FEB-2001; 2001US-0267464P.
PR	22-FEB-2001; 2001US-0271021P.
PR	14-MAR-2001; 2001US-0275946P.
PR	23-MAR-2001; 2001US-0278150P.
PR	18-APR-2001; 2001US-0284591P.
PR	23-APR-2001; 2001US-0285718P.
PR	19-JUN-2001; 2001US-0299327P.
PR	16-AUG-2001; 2001US-0312902P.
XX	
PA	(PADI/) PADIGARU M.
PA	(KEKU/) KEKUDA R.
PA	(COLM/) COLMAN S D.
PA	(SPYT/) SPYTEK K A.
PA	(BALL/) BALLINGER R A.
PA	(VERN/) VERNET C A M.
PA	(LILL/) LI L.
PA	(SHEN/) SHENOY S G.
PA	(CASM/) CASMAN S J.
XX	
PI	Padigar M, Kekuda R, Colman SD, Spytek KA, Ballinger RA;
PI	Vernet CAM, Li L, Shenoy SG, Casman SJ;
XX	
DR	WPI; 2003-755210/71.
DR	P-PSDB; AAE39674.
XX	
PT	New G-protein coupled receptor (GPCR) polypeptides and nucleic acids,
PT	useful for diagnosing, preventing or treating GPCR-associated disorders,
PT	e.g. cardiomyopathy, atherosclerosis, cancer or diabetes, and in
PT	pharmacogenomics.
XX	
XX	Claim 8; Page 40; 46pp; English.
XX	

CC The invention relates to G-protein coupled receptor (GPCR) polypeptides
 CC (designated as GPCR) and nucleic acid sequences. GPCR polypeptides,
 CC nucleic acids and antibodies are useful in diagnosing, preventing or
 CC treating GPCR-associated disorders, such as cardiomyopathy, diabetes,
 CC atherosclerosis and cancer. These may also be used for treating or
 CC preventing other diseases like developmental disorders, taste and scent
 CC detectability disorders, signal transduction pathway disorders, retinal
 CC diseases, metabolic disorders (e.g. obesity), infectious diseases (e.g.
 CC bacterial, fungal, protozoal or viral infections), bulimia, asthma,
 CC Parkinson's disease, hypertension, acute heart failure, osteoporosis,
 CC multiple sclerosis, Alzheimer's disease, stroke, graft-versus-host
 CC disease or endometriosis. GPCR nucleic acid is used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. It is also useful in gene therapy. GPCR is useful as
 CC vaccines. The present sequence is human GPCR DNA

XX Sequence 1002 BP; 242 A; 256 C; 169 G; 335 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.16e+03 Length: 1002
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 10 Gaps: 0

US-10-030-194A-5 (1-6) x AAD60343 (1-1002)

Qy 1 GlyTyr***ValGluGlu 6

Db 798 GGGTACACAGTAGAGGAA 781

RESULT 35
 AAS51422
 ID AAS51422 standard; DNA; 1035 BP.

XX AAS51422;

AC AAS51422;

DT 13-FEB-2002 (first entry)

DE Pseudomonas aeruginosa DNA for cellular proliferation protein #7.

XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW antibacterial; drug design.

XX Pseudomonas aeruginosa.

OS WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

PA Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR P-PSDB; AAU33563.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 4004; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence encodes an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 1035 BP; 180 A; 374 C; 320 G; 161 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.21e+03 Length: 1035
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x AAS51422 (1-1035)

Qy 1 GlyTyr***ValGluGlu 6

Db 964 GGCTACGGGTGAGGAA 981

RESULT 36

ACAI9457

ID ACAI9457 standard; DNA; 1035 BP.

XX ACAI9457;

DT 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #1114.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

KW drug design; gene.

XX Pseudomonas aeruginosa.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR P-PSDB; ABU15587.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

PS Claim 14; SEQ ID NO 7327; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 1035 BP; 180 A; 374 C; 320 G; 161 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.21e+03 Length: 1035
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x ACA19457 (1-1035)

Oy 1 GlyTyr***ValGluGlu 6

Db 964 GGCTACGGGTGGAGGAA 981

RESULT 37

ABD12369

ID ABD12369 standard; DNA; 1050 BP.

XX AC ABD12369;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polynucleotide #10973.

XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.

OS Pseudomonas aeruginosa.

XX US6551795-B1.

PN 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

PF

XX

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

DR P-PSDB; ABO78798.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 10973; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD1967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

SQ Sequence 1050 BP; 183 A; 378 C; 326 G; 163 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.23e+03 Length: 1050
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 11 Gaps: 0

US-10-030-194A-5 (1-6) x ABD12369 (1-1050)

Oy 1 GlyTyr***ValGluGlu 6

Db 979 GGCTACGGGTGGAGGAA 996

RESULT 38

ADC91898

ID ADC91898 standard; DNA; 1065 BP.

XX AC ADC91898;

DT 01-JAN-2004 (first entry)

DE E. faecium DNA sequence SEQ ID 1525.

XX ds; gene; urinary tract infection; bacteraemia; endocarditis; wound;
KW abdominal-pelvic infection.

OS Enterococcus faecium.

XX US6583275-B1.

PN 24-JUN-2003.

XX 30-JUN-1998; 98US-00107532.

XX 02-JUL-1997; 97US-0051571P.

PR 14-MAY-1998; 98US-0085598P.

XX

(GENO-) GENOME THERAPEUTICS CORP.
 Doucette-Stamm LA, Bush D;
 WPI; 2003-799836/75.
 P-PSDB; ADC95552.
 New isolated nucleic acid derived from Enterococcus faecium encoding an Enterococcus faecium polypeptide useful for detection, prevention and treatment of a pathological condition resulting from a bacterial infection.
 Example 1; SEQ ID NO 1525; 243pp; English.
 The invention relates to an isolated nucleic acid derived from Enterococcus faecium encoding an Enterococcus faecium polypeptide having one of 10 fully defined sequences given in the (or comprising 40 sequential nucleotides chosen from any of the nucleic acids, its complement or sequences hybridising to it). Also included are a recombinant vector comprising the nucleic acid operably linked to a transcription regulatory element, a cell comprising the vector and a single-stranded probe comprising the nucleic acid. The nucleic acids are chosen from 3654 disclosed sequences encoding 3654 disclosed proteins. The nucleic acids are useful for diagnosing pathological conditions resulting from E. faecium bacterial infection (e.g. urinary tract infection), bacteraemia, endocarditis, wounds and abdominal-pelvic infection) and for screening drugs such as agonists and antagonists. The nucleic acid is useful for recombinant production of Candida albicans - derived peptides or antisense polypeptides. Pharmaceutical compositions and vaccines containing the nucleic acid are useful for preventing or treating Enterococcus faecium infections. The present sequence represents one if the disclosed E. faecium nucleic acids.
 SQ Sequence 1065 BP; 391 A; 176 C; 222 G; 276 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.24e+03 Length: 1065
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 10 Gaps: 0
 US-10-030-194A-5 (1-6) x ADC91898 (1-1065)
 Qy 1 GlyTyr***ValGluGlu 6
 Db 361 GGCTACACGCTTGAAGAA 378
 RESULT 39
 ABN67444
 ID ABN67444 standard; DNA; 1239 BP.
 AC ABN67444;
 XX
 DT 01-JUL-2002 (first entry)
 DE Streptococcus polynucleotide SEQ ID NO 2801.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
 XX
 OS Streptococcus agalactiae.
 XX
 FN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 XX 29-OCT-2001; 2001WO-GB004789.
 PF
 XX 27-OCT-2000; 2000GB-00026333.
 PR
 XX 24-NOV-2000; 2000GB-00028727.
 PR

PR 07-MAR-2001; 2001GB-00005640.
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 DR WPI; 2002-352536/38.
 DR P-PSDB; ABP26813.
 XX
 XX New Streptococcus protein for the treatment or prevention of infection or disease caused by Streptococcus bacteria, such as meningitis, and for detecting a compound that binds to the protein.
 PT
 PT
 PS Claim 7; Page 3431; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and antiinflammatory activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and antibodies that bind (I) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. Nucleic acids encoding (I) are used to detect Streptococcus in a biological sample. (I) is used to determine whether a compound binds to (I). A composition comprising (I) or a nucleic acid encoding (I), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (I) may be used to recombinantly produce (I) and may be used in gene therapy. Antibodies to (I) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins
 XX
 SQ Sequence 1239 BP; 427 A; 179 C; 225 G; 408 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.47e+03 Length: 1239
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 6 Gaps: 0
 US-10-030-194A-5 (1-6) x ABN67444 (1-1239)
 Qy 1 GlyTyr***ValGluGlu 6
 Db 916 GGTATATAGTGTGAAGAA 933
 RESULT 40
 ADO40238
 ID ADO40238 standard; DNA; 1252 BP.
 XX
 AC ADO40238;
 XX
 DT 15-JUL-2004 (first entry)
 DE Streptococcus agalactiae NeuA gene.
 XX
 KW Group B Streptococci; GBS; capsular polysaccharide synthesis; cps;
 KW detection; NeuA; gene; ds.
 XX
 OS Streptococcus agalactiae.
 XX
 FN US2004009574-A1.
 XX
 PD 15-JAN-2004.
 XX
 XX 09-JUL-2002; 2002US-00192280.
 PF
 XX 09-JUL-2002; 2002US-00192280.
 PR

XX (DATT//) DATTAGUPTA N.
PA (SHAH//) SHAH K.
XX
XX Dattagupta N, Shah K;
PI
XX
XX WPI; 2004-090471/09.
DR
XX Novel oligonucleotide probes for detecting group B Streptococci e.g.
PT Streptococcus agalactiae in samples.
XX
XX Disclosure; Page; 28pp; English.
PS
XX The invention relates to compositions and methods for detecting group B
CC Streptococci (GBS) Streptococcus agalactiae capsular polysaccharide
CC synthesis (cps) gene in sample. The invention is useful for detecting S.
CC agalactiae in a sample. The present sequence is S. agalactiae NeuA gene.
CC This sequence is used to illustrate the method of the invention. Note:
CC This sequence is not shown in the specification, however it is
CC constructed based on the sequence provided in GenBank Accession No.
CC AF35576.
XX
SQ Sequence 1252 BP; 435 A; 176 C; 226 G; 415 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.49e+03 Length: 1252
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 12 Gaps: 0

US-10-030-194A-5 (1-6) x ADO40238 (1-1252)

Oy 1 GlyTyr***ValGluGlu 6
Db 926 GGTATAGTGTGAAGAA 943

RESULT 41
ADJ39779
ID ADJ39779 standard; cDNA; 1269 BP.
XX
AC ADJ39779;
XX
DT 06-MAY-2004 (first entry)
XX
XX Plant cDNA #779.
XX
XX Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW antifungal.
XX
XX Eukaryota.
XX
XX OS
XX US2004016025-A1.
XX
XX 22-JAN-2004.
XX
XX 26-SEP-2002; 2002US-00260238.
XX
XX 26-SEP-2001; 2001US-0325277P.
PR 26-SEP-2001; 2001US-0325448P.
PR 04-APR-2002; 2002US-0370620P.
XX
XX (BUDW//) BUDWORTH P.
PA (MOUG//) MOUGHAMER T.
PA (BRIG//) BRIGGS S P.
PA (COOP//) COOPER B.
PA (GLAZ//) GLAZEBROOK J.
PA (GOFF//) GOFF S A.

(KATA//) KATAGIRI F.
PA (KEEP//) KREPS J.
PA (PROV//) PROVART N.
PA (RICK//) RICHE D.
PA (ZHUT//) ZHU T.
XX
XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Provart N, Riche D, Zhu T;
XX
XX WPI; 2004-190374/18.
XX
XX New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT or high nutritional value.
XX
XX Claim 71; SEQ ID NO 779; 230pp; English.
PS
XX The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1269 BP; 215 A; 390 C; 389 G; 273 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 1.51e+03 Length: 1269
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 12 Gaps: 0

US-10-030-194A-5 (1-6) x ADJ39779 (1-1269)

Oy 1 GlyTyr***ValGluGlu 6
Db 313 GGATACGTCGGTGGAGGAG 330

RESULT 42
ADO80770/c
ID ADO80770 standard; DNA; 1296 BP.
XX
XX ADO80770;
AC
XX
XX 26-AUG-2004 (first entry)
XX
XX Porcine enterotoxigenic E coli resistance gene contig 79.
XX
XX ds: antibacterial; gene therapy; allele; genetic polymorphism;
KW resistance; enterotoxigenic Escherichia coli; SW2196; SW225; porcine;
KW chromosome SSC 13; antisense; iRNA; Ribozyme; gene therapy;
KW cinetoplastic DNA repair.
XX
XX Sus scrofa.
XX
XX WO2004048606-A2.
PN
XX 10-JUN-2004.
PD

XX PF 25-NOV-2003; 2003WO-DK000807.
XX PR 25-NOV-2002; 2002DK-00001819.
XX PR 25-NOV-2002; 2002US-0428738P.
XX PA (KGLV-) DEN KGL VETERINAER OG LANDBOHOUSKOLE.
XX PI Jorgensen CB, Cirera S, Archibald A, Andersson L, Fredholm M;
PI Edfors-Lilja I;
XX WPI; 2004-441182/41.
XX PT New isolated nucleic acid (NA) molecule comprising an allele of a genetic
PT polymorphism linked to resistance to enterotoxigenic Escherichia coli
PT (ETEC), useful for treating non-resistance to ETEC.
XX PS Claim 5; SEQ ID NO 151; 169pp; English.
XX CC The invention relates to an isolated nucleic acid (NA) molecule
CC comprising an allele of a genetic polymorphism linked to resistance to
CC enterotoxigenic Escherichia coli (ETEC), the genetic polymorphism being
CC located in the region between and including the markers SW2196 and SW225
CC on the porcine chromosome SSC 13. The isolated NA molecule and/or the NA
CC probe is useful as a probe for detecting a porcine allele of a genetic
CC polymorphism linked to resistant to ETEC. The NA molecules are useful as
CC antisense-NA, iRNA, Ribozyme, ETC for genetic medicine, gene therapy, and
CC sequence of the invention.
XX SQ Sequence 1296 BP; 310 A; 329 C; 313 G; 341 T; 0 U; 3 Other;

Alignment Scores: Length: 1296
Pred. No.: 1.54e+03 Matches: 5
Score: 27.00 Conserv: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.43% Gaps: 0
DB: 12

US-10-030-194A-5 (1-6) x AD080770 (1-1296)

Qy 1 GLYTYR***ValGluGlu 6
Db 900 GGATACCTAGTAGAGAG 883

RESULT 43
AAS81888
ID AAS81888 standard; cDNA; 1347 BP.
XX AC AAS81888;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #17692.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.
DR P-PSDB; ABG17701.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX PS Claim 1; SEQ ID NO 17692; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: the sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1347 BP; 274 A; 331 C; 346 G; 396 T; 0 U; 0 Other;

Alignment Scores: Length: 1347
Pred. No.: 1.61e+03 Matches: 5
Score: 27.00 Conserv: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.43% Gaps: 0
DB: 5

US-10-030-194A-5 (1-6) x AAS81888 (1-1347)

Qy 1 GLYTYR***ValGluGlu 6
Db 499 GGTACCGCTGTCGAGAA 516

RESULT 44
ADD46198
ID ADD46198 standard; DNA; 1347 BP.
XX AC ADD46198;
XX DT 29-JAN-2004 (first entry)
XX DE Rat gene U31668, SEQ ID NO 11873.
XX KW Rat; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; U31668.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat DNA (shown in Table 2 of the
CC specification) which encodes one of the polypeptides of the invention
CC which is differentially expressed during pain. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1347 BP; 417 A; 270 C; 271 G; 389 T; 0 U; 0 Other;

Alignment Scores: Length: 1347
Pred. No.: 1.61e+03 Matches: 5
Score: 27.00 Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.43% Gaps: 0
DB: 10

US-10-030-194A-5 (1-6) x ADD46198 (1-1347)

Qy 1 GlyTyr***ValGluGlu 6
Db 1174 GGATATACTGTTGAGGAA 1191

RESULT 45
ADE62406
ID ADE62406 standard; DNA; 1347 BP.
XX
AC ADE62406;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Gene U31668, SEQ ID NO 8335.
XX
KW Rat; de; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.

XX Rattus norvegicus.
OS
XX WO2003016475-A2.
PN
XX
XX 27-FEB-2003.
PD
XX
XX 14-AUG-2002; 2002WO-US025765.
PF
XX
XX 14-AUG-2001; 2001US-0312147P.
PR
XX 01-NOV-2001; 2001US-0346382P.
PR
XX 26-NOV-2001; 2001US-0333347P.
PR
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
PA
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX
XX WPI; 2003-268312/26.
DR GENBANK; U31668.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat DNA (shown in Table 2 of the
CC specification) which is differentially expressed during pain. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1347 BP; 417 A; 270 C; 271 G; 389 T; 0 U; 0 Other;

Alignment Scores: Length: 1347
Pred. No.: 1.61e+03 Matches: 5
Score: 27.00 Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.43% Gaps: 0
DB: 10

US-10-030-194A-5 (1-6) x ADE62406 (1-1347)

Qy 1 GlyTyr***ValGluGlu 6
Db 1174 GGATATACTGTTGAGGAA 1191

RESULT 46
ADC85802/c
ID ADC85802 standard; DNA; 1387 BP.

XX AC ADC85802;
XX DT 01-JAN-2004 (first entry)
XX DE Human GPCR gene SEQ ID NO:255.
XX KW ds; gene; human; GPCR;
XX KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX OS Homo sapiens.
XX FN EP1270724-A2.
XX PD 02-JAN-2003.
XX PF 18-JUN-2002; 2002EP-00013517.
XX PR 18-JUN-2001; 2001JP-00246789.
XX PA (NAAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX DR WPI; 2003-315783/31.
XX DR P-FSDB; ADC85803.
XX PT New polynucleotide, useful for preparing a composition for treating a
PT patient in need of increased or suppressed activity or expression of the
PT guanosine triphosphate-binding protein coupled receptor.
XX FS Claim 1; SEQ ID NO 255; 28pp; English.
XX CC The invention relates to a novel polynucleotide encoding a guanosine
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC the invention may have a use in gene therapy. The polynucleotide and
CC polypeptide are useful for preparing a composition for treating a patient
CC in need of increased or suppressed activity or expression of the
CC guanosine triphosphate-binding protein coupled receptor. The
CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC invention.
XX SQ Sequence 1387 BP; 370 A; 316 C; 231 G; 470 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.66e+03 Length: 1387
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 10 Gaps: 0

US-10-030-194A-5 (1-6) x ADC85802 (1-1387)

QY 1 GlyTyr***ValGluGlu 6
Db 992 GGGTACACAGTAGAGGAA 975

RESULT 47
AAF61015
ID AAF61015 standard; DNA; 1497 BP.
XX AC AAF61015;
XX DT 16-MAY-2001 (first entry)
XX DE P. putida KT2440-associated DNA ORF04209.
XX KW Transgenic plant; detection; probe; amplification; vaccine carrier;
KW microbial production strain; biological remediation; ds.
XX OS Pseudomonas putida.

XX PN DE19935088-A1.
XX PD 01-FEB-2001.
XX PF 27-JUL-1999; 99DE-01035088.
XX PR 27-JUL-1999; 99DE-01035088.
XX PA (TIGR-) TIGR INST GENOMIC RES.
PA (QUIA-) QUIAGEN GMBH.
PA (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.
PA (DKFZ-) DKFZ DEUT KREBSFORSCHUNGSZENTRUM.
PA (MEDI-) MEDIZINISCHE HOCHSCHULE HANNOVER.
XX WPI; 2001-192469/20.
XX DR New DNA sequences specific for Pseudomonas putida KT2440, useful as safe
XX genetic engineering host, allow detection in presence of other related
XX bacteria.
XX PS Claim 1a; Page 64-65; 158pp; German.
XX CC This invention describes novel DNA sequences (I) for specific detection
CC of Pseudomonas putida KT2440. The invention also describes (1)
CC recombinant expression vector containing (I); (2) prokaryotic or
CC eukaryotic cells transformed or transfected with (I) or the vector of (1)
CC; (3) production of expression products by culturing cells of (2); (4)
CC expression products, or their fragments, of (1) and synthetic proteins or
CC peptides with the same sequences (A); (5) poly- or mono-clonal antibodies
CC (Ab) that react specifically with (A); (6) hybridoma cells that produce
CC the monoclonal Ab of (5); (7) transgenic plants that contain transformed
CC or transfected cells of (2); (8) detecting KT2440 using a labeled (I) or
CC Ab as probe; and (9) DNA chips carrying one or more (I). (I), and their
CC fragments, are used as probes to detect and isolate full-length cDNAs
CC and/or to amplify such cDNAs by polymerase chain reaction, and for
CC production of transgenic plants. (I), or antibodies that recognize their
CC expression products, are used for detecting the presence of KT2440,
CC particularly in presence of other, even closely related, bacteria. KT2440
CC is one of the bacteria classified as safe, by the National Institutes of
CC Health, for genetic engineering work, e.g. as microbial production
CC strains, for biological remediation and as vaccine carriers. (I) are
CC exclusive to KT2440 with no significant homology with sequences in other
CC bacteria (specifically the closely related pathogen P. aeruginosa).
CC Compared with other 'safe' bacteria, it has greater catabolic activity
CC and better survival in, and adaptation to, the rhizosphere and soil
XX SQ Sequence 1497 BP; 275 A; 515 C; 441 G; 264 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 1.81e+03 Length: 1497
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x AAF61015 (1-1497)

QY 1 GlyTyr***ValGluGlu 6
Db 16 GGCTACAGTGTGGAAGAG 33

RESULT 48
AAS51954
ID AAS51954 standard; DNA; 1635 BP.
XX AC AAS51954;
XX DT 13-FEB-2002 (first entry)
XX DE Staphylococcus aureus DNA for cellular proliferation protein #371.
XX

KW Antisense, ds; prokaryotic cellular proliferation gene; antibiotic;
KW antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009180.
XX
XX 21-MAR-2000; 2000US-0191078P.
XX 23-MAY-2000; 2000US-0206848P.
XX 26-MAY-2000; 2000US-0207727P.
XX 23-OCT-2000; 2000US-0242578P.
XX 27-NOV-2000; 2000US-0253625P.
XX 22-DEC-2000; 2000US-0257931P.
XX 16-FEB-2001; 2001US-0269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX P-PSDB; AAU34095.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Claim 27; SEQ ID NO 4536; 51lpp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX their use in the discovery of novel antibiotics, the essential genes,
XX themselves and the encoded proteins. The prokaryotes used are Escherichia
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence encodes an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1635 BP; 538 A; 242 C; 404 G; 451 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.99e+03 Length: 1635
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 4 Gaps: 0
US-10-030-194A-5 (1-6) x AAS51954 (1-1635)
QY 1 GlyTyr***ValGluGlu 6
Db 805 GGATACCTCTGTAGAAGAA 822
RESULT 49
ACF74491
ID ACF74491 standard; DNA; 1659 BP.
XX
XX ACF74491;
XX

DT 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus DNA #2171.
XX
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX enzymatic assay; antibiotic target; gene; ds.
XX
XX Staphylococcus aureus.
XX
XX WO200294868-A2.
XX
XX 28-NOV-2002.
XX
XX 27-MAR-2002; 2002WO-IB002637.
XX
XX 27-MAR-2001; 2001GB-00007661.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Maignani V, Mora M, Scarselli M;
XX WPI; 2003-120786/11.
XX P-PSDB; ABM72931.
XX
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
XX preventing Staphylococcal infection, specifically an infection caused by
XX S. aureus, e.g. sepsis.
XX
XX Claim 6; SEQ ID NO 4341; 49pp; English.
XX
XX The invention relates to novel genes and encoded proteins from
XX Staphylococcus aureus. A composition comprising the S. aureus protein, a
XX nucleic acid encoding the protein, or an antibody to the protein, is
XX useful as a pharmaceutical, particularly as a vaccine for treating or
XX preventing infection due to Staphylococcus bacteria, specifically an
XX infection caused by S. aureus. The composition is particularly useful for
XX treating or preventing sepsis in a patient. The composition can also be
XX used for diagnostics. The protein is also used in an assay for enzymatic
XX studies and as a target for antibiotics. This sequence represents one of
XX the novel S. aureus genes of the invention
XX
XX Sequence 1659 BP; 546 A; 244 C; 413 G; 456 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2.02e+03 Length: 1659
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0
US-10-030-194A-5 (1-6) x ACF74491 (1-1659)
QY 1 GlyTyr***ValGluGlu 6
Db 808 GGATACCTCTGTAGAAGAA 825
RESULT 50
AAS54463
ID AAS54463 standard; DNA; 1662 BP.
XX
XX AAS54463;
XX
XX 13-FEB-2002 (first entry)
XX
XX Staphylococcus aureus DNA for cellular proliferation protein #775.
XX
XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
XX antibacterial; drug design.
XX
XX Staphylococcus aureus.
XX
XX WO200170955-A2.
XX

```
XX PD 27-SEP-2001.
XX XX
XX PF 21-MAR-2001; 2001WO-US009180.
XX XX
XX PR 21-MAR-2000; 2000US-0191078P.
XX PR 23-MAY-2000; 2000US-0206848P.
XX PR 26-MAY-2000; 2000US-0207727P.
XX PR 23-OCT-2000; 2000US-0242578P.
XX PR 27-NOV-2000; 2000US-0253625P.
XX PR 22-DEC-2000; 2000US-0257931P.
XX PR 16-FEB-2001; 2001US-0269308P.
XX XX
XX PA (ELIT-) ELITRA PHARM INC.
XX FI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX XX
XX DR WPI; 2001-611495/70.
XX DR P-PSDB; AAU36604.
XX XX
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX PS Claim 27; SEQ ID NO 8100; 511pp; English.
XX XX
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes
XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia
XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence encodes an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1662 BP; 550 A; 243 C; 411 G; 458 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.03e+03 Length: 1662
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x AAS54463 (1-1662)
Qy 1 GlyTyr***ValGluGlu 6
| | | | | | | |
Db 808 GGATACTCTGTAGAAGAA 825

RESULT 51
ACA20007
ID ACA20007 standard; DNA; 1662 BP.
XX AC
XX AC ACA20007;
XX XX
XX DT 19-JUN-2003 (first entry)
XX XX
XX DE Prokaryotic essential gene #1664.
XX XX
XX KW Antisense; db; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
```

```
XX OS Staphylococcus aureus.
XX XX
XX PN WO200277183-A2.
XX XX
XX PD 03-OCT-2002.
XX XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX XX
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX XX
XX PA (ELIT-) ELITRA PHARM INC.
XX XX
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX XX
XX DR WPI; 2003-029926/02.
XX DR P-PSDB; ABU16137.
XX XX
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 14; SEQ ID NO 7877; 1766pp; English.
XX XX
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
XX CC prokaryotic essential genes. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1662 BP; 550 A; 243 C; 412 G; 457 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.03e+03 Length: 1662
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x ACA20007 (1-1662)
```

QY 1 GlyTyr***ValGluGlu 6
Db 808 GGATACCTCTAGAGAA 825
RESULT 52
AAS81886/c
ID AAS81886 standard; cDNA; 1857 BP.
XX AAS81886;
AC
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #17690.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS
XX WO200175067-A2.
FN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR
XX 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG17699.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 17690; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1857 BP; 421 A; 515 C; 513 G; 408 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,29e+03 Length: 1857
Score: 27.00 Matches: 5
Percent Similarity: 83.3% Conservative: 0
Best Local Similarity: 83.3% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x AAS81886 (1-1857)
QY 1 GlyTyr***ValGluGlu 6
Db 1767 GGTACCGCTGTCGAGGAA 1750
RESULT 53
AAF26350/c
ID AAF26350 standard; DNA; 1896 BP.
XX
AC AAF26350;
XX
DT 02-MAY-2001 (first entry)
DE P. putida oxygenase encoding DNA ORF04210.
XX
KW Oxygenase; transgenic plant; detection; amino acid production;
KW vitamin production; steroid production; dialcohol production;
KW oxidized aromatic compound production; dialdehyde production;
KW optically active epoxide production; dicarboxylic acid production;
KW bioconversion; ORF04210; ds.
XX
OS Pseudomonas putida.
XX
FN WO200107629-A2.
XX
PD 01-FEB-2001.
XX
PF 27-JUL-2000; 2000WO-EP007244.
XX
PR 27-JUL-1999; 99DE-01035087.
XX
XX (TIGR-) TIGR INST GENOMIC RES.
PA (QUIA-) QUIAGEN GMBH.
PA (GSFB) GBS BIOEKOLOGISCHE FORSCHUNG MBH.
PA (DKFZ-) DKFZ DEUT KREBSFORSCHUNGSZENTRUM.
PA (MEDI-) MEDIZINISCHE HOCHSCHULE HANNOVER.
XX
PI Fraser CM, Venter C, Tuemmler B, Hoheisel J, Duesterhoeft A;
PI Hilbert H, Timmis KN, Moore E, Straetz M, Heim S;
XX
XX WPI; 2001-168556/17.
XX
XX New DNA encoding a bacterial oxygenase, useful e.g. for production of
PT amino acids or vitamins, also derived transgenic plants.
XX
PS Claim 3a; Page 33-34; 47pp; German.
XX
XX This invention describes novel DNA sequences (I) that express products
CC having the biological function of oxygenases. The invention also
CC describes (1) recombinant expression vectors containing (I); (2)
CC prokaryotic and eukaryotic cells transformed or transfected with (1) or
CC the vector of (1); (3) production of oxygenases by culturing cells of (2)
CC ; (4) (partial) expression products (II) of (1), and synthetic proteins
CC or peptides with the same sequences; (5) mono- or polyclonal antibodies
CC (Ab) specific for (II); (6) hybridoma cells that produce monoclonal Ab;
CC and (7) transgenic plants that contain cells of (2). (I), and their
CC fragments, are useful for the following: (i) expression of recombinant
CC oxygenases; (ii) useful as probes and primers for detection, isolation
CC and amplification of full-length cDNA sequences; and (iii) used to
CC produce transgenic plants. Expression products are useful for production
CC of amino acids, vitamins, steroids, oxidized aromatic compounds,
CC optically active epoxides, dialcohols, dialdehydes and dicarboxylic
CC acids. The transformed cells that express them are used for bioconversion
CC of aromatic and aliphatic compounds. The oxygenases have very weak
CC substrate specificity so can be used for highly stereo- or regio-specific
CC oxidations of a wide range of substrates, reactions that are difficult or
CC impossible to do with conventional oxidizing agents
XX
SQ Sequence 1896 BP; 340 A; 522 C; 643 G; 391 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.34e+03 Length: 1896
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x AAF26350 (1-1896)

Qy 1 GlyTyr***ValGluGlu 6
 Db 51 GGCTACAGTGTGGAAGAG 34

RESULT 54

ID ACC61239 standard; DNA; 2000 BP.

XX AC ACC61239;

XX 20-JUN-2003 (first entry)

XX Gene sequence #SEQ ID 1260.

XX Multiprotein complex; eukaryote; drug target; diagnosis; gene; ds.

XX Saccharomyces cerevisiae.

XX EPI258494-A1.

XX 20-NOV-2002.

XX 20-DEC-2001; 2001EP-00130253.

XX 15-MAY-2001; 2001EP-00111774.

XX (CELL-) CELLZONE AG.

PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;

PI Marzioch M, Schultz JD, Superti-Furga GD;

XX WPI; 2003-250078/25.

DR P-PSDB; ABR53197.

XX New isolated protein complexes useful for diagnosing a disease or

PT disorder, or as a target for an active agent of a pharmaceutical,

PT preferably a drug target in the treatment or prevention of disease or

PT disorder.

PS Disclosure; SEQ ID NO 1260; 17pp + Sequence Listing; English.

XX The invention relates to multiprotein complexes from eukaryotes. Proteins
 CC of the invention and DNA sequences encoding them are given in records
 CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
 CC obtainable by using a protein as a bait and isolating the set of proteins
 CC which is attached thereto from cells. Such protein complexes may comprise
 CC up to 30 distinct proteins. Protein complexes of the invention are useful
 CC for diagnosing a disease or disorder, or as a target for an active agent
 CC of a pharmaceutical, preferably a drug target in the treatment or
 CC prevention of a disease or disorder. Note: The sequence data for this
 CC patent is not represented in the printed specification, but is based on
 CC sequence information supplied by the European Patent Office. The complete
 CC document is available on CD-ROM

XX Sequence 2000 BP; 671 A; 327 C; 360 G; 642 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.48e+03 Length: 2000
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 10 Gaps: 0

US-10-030-194A-5 (1-6) x ACC61239 (1-2000)

Qy 1 GlyTyr***ValGluGlu 6
 Db 1693 GGTTATACGGTAGAGAA 1710

RESULT 55

ID ADK63243 standard; DNA; 2000 BP.

XX AC ADK63243;

XX 06-MAY-2004 (first entry)

XX Disease treating protein complex-derived gene #747.

XX protein complex; drug target; diagnosis; gene; ds.

XX Unidentified.

XX EPI338608-A2.

XX 27-AUG-2003.

XX 20-DEC-2002; 2002EP-00102902.

XX 20-DEC-2001; 2001EP-00130253.

XX (CELL-) CELLZONE AG.

PI Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;

PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;

PI Michon A, Leutwein C, Rick J;

XX WPI; 2003-638460/61.

DR P-PSDB; ADK63242.

XX New proteins and protein complexes from eukaryotes, useful as targets in
 PT drug screening, or in diagnosing or screening for the presence of a
 PT disease or disorder, or a predisposition for developing a disease or
 PT disorder in a subject.

PS Disclosure; SEQ ID NO 1494; 13pp; English.

XX The invention relates to novel protein complexes comprising a first and a
 CC second protein, or its derivative, fragment, homologue or variant. The
 CC proteins are selected from given protein complexes, which are not defined
 CC in the specification. The variants are encoded by nucleic acids that
 CC hybridize to the nucleic acids encoding the proteins under low stringency
 CC conditions. The protein complexes are useful as targets for an active
 CC agent of a pharmaceutical. These protein complexes are particularly
 CC useful as drugs targets for the treatment or preventing of a disease or
 CC disorder. The complexes and methods above are useful in diagnosing or
 CC screening for the presence of a disease or disorder or a predisposition
 CC for developing a disease or disorder in a subject. These are also useful
 CC in screening for a drug for treatment or prevention of a disease or
 CC disorder. The molecule that modulates the amount, activity or protein
 CC components of the complex is useful for the manufacture of a medicament
 CC for the treatment or prevention of a disease or disorder. This sequence
 CC corresponds to a gene of the invention. (Note: the sequence data for this
 CC patent did not form part of the printed specification but was obtained
 CC from the EPO in electronic format).

XX Sequence 2000 BP; 671 A; 327 C; 360 G; 642 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.48e+03 Length: 2000
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 10 Gaps: 0

US-10-030-194A-5 (1-6) x ADK63243 (1-2000)

QY 1 GlyTyr***ValGluGlu 6
||||| |||||
Db 1693 GGTATACGTTAGAGAA 1710

RESULT 56

ADP38040/c
ID ADF38040 standard; cDNA; 2036 BP.

XX AC ADF38040;

XX DT 12-FEB-2004 (first entry)

XX DE Synchronised tobacco BY2 cDNA sequence SEQ ID NO:99.

XX KW identification; validation; plant; agrochemical; gene; ss.

XX OS Nicotiana tabacum.

XX PN WO2003085115-A2.

XX PD 16-OCT-2003.

XX PF 08-APR-2003; 2003WO-EP003703.

XX PR 10-APR-2002; 2002EP-00447062.

XX PR 15-JUL-2002; 2002US-0396124P.

XX PA (CROP-) CROPDISEIGN NV.

XX PI Inze D, Broekaert W;

XX DR WPI; 2003-804308/75.

XX PT Identifying and validating plant genes or proteins as targets for agrochemicals, useful for producing agrochemical-resistant plants, comprises determining and down regulating the gene or protein expression profiles of a plant.

XX PS Claim 12; SEQ ID NO 99; 183pp; English.

XX CC The present invention describes a method for identifying and validating plant genes/proteins as targets for agrochemicals comprising determining the gene or protein expression profiles of a plant and downregulating the expression of the gene or protein in the plant or plant cell. Also described: (1) methods for screening candidate agrochemical compounds, comprising the use of the above method or the use of any of the 785 fully defined nucleotide sequences (ADP37942 to ADF38726) or protein sequences, or their homologues, functional fragments or derivatives; (2) a method for producing an agrochemical resistant plant, comprising the use of the above-mentioned nucleotide or protein sequences; (3) an isolated nucleic acid that is identified by any of the above methods or that comprises at least a part of a nucleic acid sequence chosen from any of the 785 nucleotide sequences given in the specification; (4) a plant tolerant to an agrochemical, in which the expression level of one or more of the nucleic acid sequences given in the specification is modulated; and (5) a harvestable part of the plant described above. The method is useful in identifying and validating plant targets for agrochemicals or in producing an agrochemical resistant plant. The nucleic acid or protein can be used as a target for an agrochemical compound, particularly herbicide. The present sequence represents a synchronised tobacco BY2 cDNA nucleotide sequence which is used in the exemplification of the present invention.

XX SQ Sequence 2036 BP; 539 A; 399 C; 439 G; 659 T; 0 U; 0 Other;

Alignment Scores:

Score: No.:	2.53e+03	Length:	2036
Pred. No.:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0

DB: 10 Gaps: 0

US-10-030-194A-5 (1-6) x ADF38040 (1-2036)

QY 1 GlyTyr***ValGluGlu 6
||||| |||||

Db 1929 GGTATACAGTAGAGAA 1912

RESULT 57

AAAD07614

ID AAD07614 standard; cDNA; 2094 BP.

XX AC AAD07614;

XX DT 10-AUG-2001 (first entry)

XX DE Human secreted protein-encoding gene 9 cDNA clone HMWEJ52, SEQ ID NO:54.

XX KW Human; secreted protein; proliferative disorder; cancer; tumour; foetal abnormality; developmental abnormality; haematopoietic disorder; immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis; inflammation; allergy; neurological disorder; Alzheimer's disease; Parkinson's disease; cognitive disorder; schizophrenia; asthma; skin disorder; psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder; angiogenic disorder; kidney disorder; gastrointestinal disorder; pregnancy-related disorder; gene therapy; endocrine disorder; infection; wound healing; vulnerability; cell culture; chemotaxis; food additive; binding partner identification; chromosome 15; ss.

XX OS Homo sapiens.

XX FT Key Location/Qualifiers

XX CDS 336..491

XX FT /*tag= a

XX FT /product= "Human secreted protein precursor"

XX FT sig_peptide 336..407

XX FT /*tag= b

XX FT mat_peptide 408..488

XX FT /*tag= c

XX FT /product= "Mature human secreted protein"

XX PN WO200132676-A1.

XX PD 10-MAY-2001.

XX PF 25-OCT-2000; 2000WO-US029365.

XX PR 29-OCT-1999; 99US-0162237P.

XX PR 21-JUL-2000; 2000US-0219666P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsoulis GA, Shi Y, Olsen HS, Soppet DR;

XX WPI; 2001-328773/34.

XX DR P-PSDB; AAE03095.

XX PT Nucleic acids encoding 25 human secreted polypeptides, useful for preventing, diagnosing and/or treating e.g. Gaucher's disease, PT Alzheimer's disease, Scimitar syndrome, Creutzfeldt-Jacob disease, PT diabetes mellitus and multiple sclerosis.

XX PS Claim 1; Page 436-437; 546pp; English.

XX CC AAD07571-AAD07645 represent cDNAs corresponding to 25 human secreted protein genes, and AAE03052-AAE03126 represent the proteins they encode. CC AAE03127-AAE03150 represent human secreted protein fragments. The genes CC and their corresponding secreted proteins are useful for preventing, CC treating or ameliorating medical conditions, e.g., by protein or gene CC therapy. Pathological conditions can be diagnosed by determining the CC amount of the new protein in a sample or by determining the presence of CC mutations in the new genes. Specific uses are described for each of the

25 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis or treatment of proliferative disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplant, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein-encoding cDNA of the invention

Sequence 2094 BP; 515 A; 479 C; 521 G; 579 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.61e+03 Length: 2094
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservativity: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x AAD07614 (1-2094)

Qy 1 GlyTyr***ValGluGlu 6
 |||||
 Db 229 GGATATGCAGTAGAGGAA 246

RESULT 58

ADM03396/c
 ID ADM03396 standard; cDNA; 2104 BP.

AC ADM03396;

DT 20-MAY-2004 (first entry)

XX Human cDNA of the invention SEQ ID NO:2081.

DE ss; Gene; human; Gene therapy; diagnostic marker; pharmaceutical.

KW Homo sapiens.

OS EP1347046-A1.

PN 24-SEP-2003.

XX 12-APR-2002; 2002EP-00008400.

PF 22-MAR-2002; 2002JP-00137785.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-723558/69.

DR P-PSDB; ADM05839.

XX New polynucleotides and polypeptides are useful in gene therapy, for
 PT developing a diagnostic marker or medicines for regulating their
 PT expression and activity, or as a target of gene therapy.

PS Claim 1; SEQ ID NO 2081; 305pp; English.
 XX The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC cDNA sequence of the invention.

XX Sequence 2104 BP; 493 A; 588 C; 508 G; 515 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.62e+03 Length: 2104
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservativity: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 11 Gaps: 0

US-10-030-194A-5 (1-6) x ADM03396 (1-2104)

Qy 1 GlyTyr***ValGluGlu 6
 |||||
 Db 2058 GGATATGCAGTAGAGGAA 2041

RESULT 59

AAX27286

ID AAX27286 standard; DNA; 2289 BP.

XX AAX27286;

DT 02-JUN-1999 (first entry)

XX Desulfurococcus M11 TL DNA polymerase 29PY1 coding sequence.

DE DNA polymerase; thermophilic bacteria; DNA synthesis; ss.

OS Desulfurococcus sp.

PN WO9907837-A1.

PD 18-FEB-1999.

XX 06-AUG-1998; 98WO-US017152.

PR 06-AUG-1997; 97US-00907166.

XX (DIVE-) DIVERSA INC.

XX Callen W, Mathur EJ;

XX WPI; 1999-180490/15.

DR P-PSDB; AAY00939.

XX DNA polymerases from extremely thermophilic bacteria - useful for DNA
 PT synthesis.

XX Claim 3; Fig 5; 72pp; English.

XX This sequence encodes a DNA polymerase of the invention, that was
 CC isolated from a thermophilic bacteria. The polymerases are used in DNA
 CC synthesis and as immunogens to raise antibodies (useful for affinity
 CC purification and to screen for related enzymes). Fragments of the DNA
 CC encoding the polymerases are used as probes to isolate related or full-
 CC length sequences and to produce the recombinant polymerases. The
 CC polymerases catalyse DNA synthesis by the addition of deoxynucleotides to
 CC the 3' end of a polynucleotide chain, using a complementary
 CC polynucleotide strand as a template. The polymerases have optimum
 CC activity at over 60 deg. C and can renature and regain activity after

CC exposure to temperatures above 70 deg. C
SQ Sequence 2289 BP; 648 A; 483 C; 631 G; 526 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2,88e+03 Length: 2289
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-5 (1-6) x AAX27286 (1-2289)

QY 1 GlyTyr***ValGluGlu 6
|||||
DB 2065 GGGTACAGCGTGGAGGAG 2082

RESULT 60

AAD35185
ID AAD35185 standard; DNA; 2289 BP.

AC AAD35185;

DT 25-JUL-2002 (first entry)

DE Desulfurococcus sp. DNA polymerase encoding DNA.

KW DNA polymerase; thermostable; enzyme; gene; ds.

OS Desulfurococcus sp.

FH Key Location/Qualifiers
FT CDS 1..2289
/*tag= a

FT /product= "Desulfurococcus sp. DNA polymerase"
FT /transl_except= (pos:1801..1803, aa:Xaa)
FT /note= "Xaa = Ala, Pro"

PN WO200220735-A2.

XX 14-MAR-2002.

PF 06-SEP-2001; 2001WO-US028007.

PR 06-SEP-2000; 2000US-00656309.

XX (DIVE-) DIVERSA CORP.

PI Callen W, Mathur EJ, Short JM;

DR WPI; 2002-362247/39.

DR P-PSDB; AAE22116.

XX New thermostable polymerase useful for sequencing DNA, amplifying double
PT stranded DNA, or incorporating a non-natural nucleotide or a nucleotide
PT analog into a DNA molecule.

PS Claim 1; Page 146-147; 161pp; English.

CC The invention relates to thermostable DNA polymerases having high
CC temperature polymerase activity, such as those derived from Pyrobolus
CC fumaria and nucleic acid molecules encoding such polymerases. Polymerases
CC are useful for catalyzing the formation or repair of a nucleic acid
CC sequence and for modifying small molecules. They are also useful for
CC sequencing DNA molecules, for preparing cDNA from mRNA, for amplifying
CC double stranded DNA molecules and for incorporating non-natural
CC nucleotides or nucleotide analogues into a DNA molecule. The present
CC sequence is Desulfurococcus sp. DNA polymerase encoding DNA

SQ Sequence 2289 BP; 648 A; 483 C; 631 G; 526 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 2,88e+03 Length: 2289
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x AAD35185 (1-2289)

QY 1 GlyTyr***ValGluGlu 6
|||||
DB 2065 GGGTACAGCGTGGAGGAG 2082

RESULT 61

ABX14889
ID ABX14889 standard; DNA; 2289 BP.

AC ABX14889;

XX 08-APR-2003 (first entry)

XX DNA encoding Desulfurococcus thermostable DNA polymerase protein.

KW Gene; ds; thermostable; DNA polymerase; DNA repair;

KW polymorphism identification.

OS Desulfurococcus sp.

FH Key Location/Qualifiers
FT CDS 1..2289
/*tag= a

FT /product= "Desulfurococcus DNA polymerase"
FT US2002132243-A1.

XX 19-SEP-2002.

XX 06-SEP-2001; 2001US-00948369.

XX 06-AUG-1997; 97US-00907166.

PR 07-SEP-1999; 99US-00391340.

PR 06-SEP-2000; 2000US-00656309.

XX (CALL/) CALLEN W.

PA (MATH/) MATHUR E J.

PA (SHOR/) SHORT J.

XX Callen W, Mathur EJ, Short J;

XX WPI; 2003-182285/18.

DR P-PSDB; ABG73160.

XX Novel DNA polymerases having increased activity and stability at
PT increased pH and temperature, useful for DNA sequencing, amplification
PT and incorporating non-natural nucleotides or nucleotide analogs.

PS Claim 1; Page 45-46; 81pp; English.

CC This invention relates to a novel purified polypeptide derived from
CC Pyrobolus fumari which has thermostable DNA polymerase activity. The
CC protein of the invention is useful for catalyzing the formation or repair
CC of a nucleic acid sequence, for comparing a sequence to a reference
CC sequence and for identifying polymorphisms. The nucleic acid sequence of
CC the invention is useful for preparing cDNA from mRNA, and in an
CC amplification process of a double-stranded DNA molecule. The nucleotide
CC sequence is also useful for incorporating non-natural nucleotides or its
CC analogues into a DNA molecule, by contacting a polypeptide encoded by the
CC nucleic acid with a DNA template in a PCR amplification reaction. The
CC nucleotides which may be used for this are inosine, 2-aminopurine or 5-
CC methylcytosine. A nucleic acid probe derived from the nucleic acid
CC sequence of the invention is useful in chromosome walking procedures to
CC identify clones containing genomic sequences located adjacent to a
CC sequence of the gene encoding the DNA polymerases of the invention. Such

CC methods allow the isolation of genes which encode additional proteins
 CC from the host organisms. The polymerases of the invention have increased
 CC activity and stability at increased pH and temperature, and high
 CC processivity. The present sequence represents a DNA sequence encoding a
 CC Desulphurococcus thermostable DNA polymerase protein of the invention
 XX
 SQ Sequence 2289 BP; 648 A; 483 C; 631 G; 526 T; 0 U; 1 Other;

Alignment Scores: Length: 2289
 Pred. No.: 2.88e+03
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x ABX14889 (1-2289)

QY 1 GlyTyr***ValGluGlu 6
 DB 2065 GGGTACACGGTGGAGGAG 2082

RESULT 62
 ABD12667/c
 ID ABD12667 standard; DNA; 2547 BP.

AC ABD12667;

XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polynucleotide #11271.

XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
 KW antibacterial.

OS Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

DR P-PsDB; ABO79096.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 11271; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABD01397-
 CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
 CC The sequence data for this patent did not form part of the printed

CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 2547 BP; 452 A; 778 C; 942 G; 375 T; 0 U; 0 Other;

Alignment Scores: Length: 2547
 Pred. No.: 3.23e+03
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 11 Gaps: 0

US-10-030-194A-5 (1-6) x ABD12667 (1-2547)

QY 1 GlyTyr***ValGluGlu 6
 DB 42 GGCTACCGGTGGAGGAA 25

RESULT 63

AAS62441
 ID AAS62441 standard; cDNA; 2682 BP.

XX AAS62441;

XX 14-FEB-2002 (first entry)

XX cDNA sequence #228 encoding novel human secreted protein.

XX Human secreted protein; hyperproliferative disorder; autoimmune disorder;
 KW immune deficiency disorder; blood disorder; inflammatory disorder;
 KW infectious disorder; gene therapy; antimicrobial; hepatotropic;
 KW immunosuppressive; antirheumatic; ss.

OS Homo sapiens.

XX WO200177291-A2.

XX 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US010485.

PR 06-APR-2000; 2000US-0195604P.

PA (GEMY) GENETICS INST INC.

XX Wong GG, Clark HF, Fichtel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Gulukota K, Graham JR;

XX WPI; 2002-010900/01.

XX New polynucleotides encoding secreted proteins useful for treating e.g.
 PT asthma, HIV and Crohn's disease.

XX Claim 1; Page 194; 391pp; English.

XX The present invention relates to the isolation of novel cDNA sequences
 CC which encode human secreted proteins. The cDNA sequences have been
 CC derived from a variety of human tissues. The invention also provides a
 CC method for producing proteins from these polynucleotide sequences. The
 CC proteins are useful for identifying compounds that modulate their
 CC activity and production, and the cell is also useful for identifying
 CC compounds that modulate expression of the polynucleotide sequences
 CC encoding the secreted proteins. The sequences of the invention are useful
 CC for treating diseases such as hyperproliferative disorders (e.g. cancer),
 CC immune deficiency disorders (e.g. severe combined immunodeficiency
 CC (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
 CC (e.g. thrombocytopenia), inflammatory disorders (e.g. arthritis) and
 CC infectious disorders (e.g. hepatitis). The polynucleotide sequences of
 CC the invention are also useful in gene therapy. AAS62214-AAS62838
 CC represent the cDNA sequences of the invention that encode for novel human
 CC secreted proteins

Alignment Scores:

Percent Similarity: 83.33%
Best Local Similarity: 83.33%

Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x ABV24673 (1-3645)

Qy 1 GlyTyr***ValGluGlu 6
DB 1505 GGATATGCAGTAGAGGAA 1522

RESULT 66
ABV28517
ID ABV28517 standard; cDNA; 3645 BP.
XX
AC ABV28517;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 28508.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
XX
PR 16-MAR-2000; 2000US-0189862P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 5959-5960; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 3645 BP; 876 A; 963 C; 984 G; 822 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.79e+03 Length: 3645
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x ABV28517 (1-3645)

Qy 1 GlyTyr***ValGluGlu 6
DB 1505 GGATATGCAGTAGAGGAA 1522

RESULT 67
ABV22692
ID ABV22692 standard; cDNA; 3645 BP.
XX
AC ABV22692;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 22683.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
XX
PR 16-MAR-2000; 2000US-0189862P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3985-3986; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 3645 BP; 876 A; 963 C; 984 G; 822 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.79e+03 Length: 3645
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x ABV22692 (1-3645)

Qy 1 GlyTyr***ValGluGlu 6
DB 1505 GGATATGCAGTAGAGGAA 1522

```
RESULT 68
ABV29547
ID ABV29547 standard; cDNA; 3645 BP.
XX
XX
AC ABV29547;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 29538.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
XX 25-MAY-2000; 2000US-0207454P.
XX
XX 09-JUN-2000; 2000US-0211314P.
XX
XX 18-JUL-2000; 2000US-0219007P.
XX
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 6330-6331; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 3645 BP; 876 A; 963 C; 984 G; 822 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4.79e+03 Length: 3645
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0
US-10-030-194A-5 (1-6) x ABV29547 (1-3645)
Qy 1 GlyTyr***ValGluGlu 6
Db 1505 GGATATGCAGTAGAGGAA 1522
RESULT 69
ABL58963
ID ABL58963 standard; cDNA; 3962 BP.
XX
XX
AC ABL58963;
XX
XX 19-JUL-2002 (first entry)
XX
XX Human tumour marker cDNA se33-1.
XX
XX Human; tumour; cytostatic; cutaneous T cell lymphoma; CTL; vaccine;
XX antigen-presenting cell; tumour-specific T cell; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200238803-A2.
XX
XX 16-MAY-2002.
XX
XX 08-NOV-2001; 2001WO-DE004229.
XX
XX 08-NOV-2000; 2000DE-01055285.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Eichmueller S, Schadendorf D, Usener D;
XX
XX WPI; 2002-425959/45.
XX
XX P-PSDB; ABB77436.
XX
XX Composition containing tumor-associated nucleic acid, useful for
XX diagnosis and treatment of tumors, especially cutaneous T cell lymphoma,
XX also derived proteins and antibodies.
XX
XX Claim 1; Fig 13; 84pp; German.
XX
XX The invention relates to a diagnostic composition containing at least one
XX of 23 nucleotide sequences (I, ABL58901-ABL58950) with altered expression
XX associated with tumors. (I), including antisense sequences and
XX ribozymes, also proteins (II, ABB77424-ABB77445) encoded by them and
XX antibodies specific for (II), are useful for diagnosis, monitoring and
XX treatment of tumors, especially cutaneous T cell lymphoma (CTCL). (II)
XX and antibodies to (II) are useful for vaccination. (II) can also be used
XX to prepare pre-loaded antigen-presenting cells or tumour-specific T cells
XX
XX Sequence 3962 BP; 1455 A; 627 C; 867 G; 1013 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 5.25e+03 Length: 3962
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0
US-10-030-194A-5 (1-6) x ABL58963 (1-3962)
Qy 1 GlyTyr***ValGluGlu 6
Db 502 GGATATTCGTAGAAGAA 519
RESULT 70
ACA21975
ID ACA21975 standard; DNA; 4217 BP.
XX
XX ACA21975;
XX
XX 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene #3632.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
XX Bacillus anthracis.
XX
```

PN WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU18105.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 14; SEQ ID NO 9845; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway;
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 4217 BP; 1515 A; 735 C; 875 G; 1092 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,62e+03 Length: 4217
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 8 Indels: 0
 DB: Gaps: 0
 US-10-030-194A-5 (1-6) x ACA21975 (1-4217)
 Oy 1 GlyTyr***ValGluGlu 6
 Db 853 GGTACGCAGTAGAGAA 870

RESULT 71
 AAS81892/c
 ID AAS81892 standard; cDNA; 5728 BP.
 XX AAS81892;
 AC AAS81892;
 XX 13-FEB-2002 (first entry)
 DT 13-FEB-2002 (first entry)
 XX DNA encoding novel human diagnostic protein #17696.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 KW Homo sapiens.
 OS WO200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US008631.
 PF 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABG17705.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 1; SEQ ID NO 17696; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activities. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 5728 BP; 1333 A; 1511 C; 1593 G; 1290 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 7,86e+03 Length: 5728
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: Gaps: 0
 US-10-030-194A-5 (1-6) x AAS81892 (1-5728)

QY 1 GlyTyr***ValGluGlu 6
Db 3721 GGTACGCTGTCAGGAA 3704

RESULT 72
ABL33576
ID ABL33576 standard; DNA; 5925 BP.
XX
AC ABL33576;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1549.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytosatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriasis;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
XX Homo sapiens.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 1549; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 5925 BP; 1510 A; 105 C; 1246 G; 3064 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.16e+03 Length: 5925
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x ABL33576 (1-5925)

QY 1 GlyTyr***ValGluGlu 6
Db 2028 GGATATAGTGTGAGGAA 2045

RESULT 73
ACF34525

ID ACF34525 standard; DNA; 6570 BP.
XX
AC ACF34525;
XX
DT 15-OCT-2003 (first entry)
XX
DE Gene encoding angiogenesis protein BNO159.
XX
KW Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antipsoriasis; antiarteriosclerotic; cardiatic; vasotropic; angiogenesis;
KW gene therapy; vasculature; cancer; rheumatoid arthritis; psoriasis;
KW diabetic retinopathy; cardiovascular disease; atherosclerosis;
KW ischemic limb disease; coronary artery disease; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003027285-A1.
XX
PD 03-APR-2003.
XX
PF 19-SEP-2002; 2002WO-AU001282.
XX
PR 27-SEP-2001; 2001AU-00007973.
PR 27-SEP-2001; 2001AU-00007974.
PR 11-OCT-2001; 2001AU-00008210.
PR 29-OCT-2001; 2001AU-00008532.
PR 13-NOV-2001; 2001AU-00008838.
PR 28-AUG-2002; 2002AU-00951032.
XX
PA (BION-) BIONOMICS LTD.
XX
PI Gamble JR, Hahn CN, Vadas MA;
XX
DR WPI; 2003-354655/33.
DR P-PSDB; ABR64250.
XX
PT New angiogenic genes and polypeptides, useful for diagnosing,
PT prognosticating or treating an angiogenesis-related disorder, e.g.
PT cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or
PT cardiovascular diseases.
XX
PS Claim 2; SEQ ID NO 80; 90pp; English.
XX
CC The invention relates to the isolation of novel genes (ACF34446-ACF34559)
CC encoding proteins (ABR64180-ABR64281) involved in the process of
CC angiogenesis. The nucleic acid molecules are useful in identifying and/or
CC obtaining full-length human genes involved in an angiogenic process. The
CC nucleic acid molecule, polypeptides or complexes encoded, cells or
CC genetically modified non-human animals derived from these are useful for
CC the screening of candidate pharmaceutical compounds used in treating
CC angiogenesis-related disorders. They are also useful for diagnosing,
CC prognosticating or treating an angiogenesis-related disorder, which
CC involves uncontrolled or enhanced angiogenesis or is a disorder in which
CC a decreased vasculature is of benefit (e.g. cancer, rheumatoid arthritis,
CC diabetic retinopathy, psoriasis or cardiovascular diseases such as
CC atherosclerosis), or involves inappropriately arrested or decreased
CC angiogenesis or is a disorder in which an expanding vasculature is of
CC benefit (e.g. ischemic limb disease or coronary artery disease). The
CC modulator of expression or activity of the polypeptide encoded by the
CC nucleic acid sequence is useful for manufacturing a medicament for the
CC treatment of an angiogenesis-related disorder. This sequence corresponds
XX to the gene encoding one of the novel angiogenic protein
SQ Sequence 6570 BP; 2337 A; 1165 C; 1412 G; 1656 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.14e+03 Length: 6570
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x ACF34525 (1-6570)
 QY 1 GlyTyr***ValGluGlu 6
 Db 3060 GGATATTCTGTAGAAGAA 3077
 RESULT 74
 ABL69964
 ID ABL69964 standard; DNA; 6571 BP.
 AC ABL69964;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Pancreas cancer related gene sequence SEQ ID NO:8301.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200194629-A2.
 FN
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US010838.
 XX
 PR 05-JUN-2000; 2000US-0209473P.
 PR 05-JUN-2000; 2000US-0209531P.
 PR 18-SEP-2000; 2000US-0233133P.
 PR 18-SEP-2000; 2000US-0233617P.
 PR 20-SEP-2000; 2000US-0234009P.
 PR 20-SEP-2000; 2000US-0234034P.
 PR 20-SEP-2000; 2000US-0234052P.
 PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234923P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237318P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 01-NOV-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 PA (AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 FS Claim 1; SEQ ID NO 8301; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour
 XX
 SQ Sequence 6571 BP; 2335 A; 1164 C; 1415 G; 1657 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 9.14e+03 Length: 6571
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 6 Gaps: 0
 US-10-030-194A-5 (1-6) x ABL69964 (1-6571)
 QY 1 GlyTyr***ValGluGlu 6
 Db 3061 GGATATTCTGTAGAAGAA 3078
 RESULT 75
 ABR83708
 ID ABR83708 standard; cDNA; 6571 BP.
 AC ABR83708;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Human cDNA differentially expressed in granulocytic cells #279.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX
 OS Homo sapiens.
 XX
 FN WO200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001WO-US030821.
 XX

```

PR 03-OCT-2000; 2000US-0237189P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
XX WPI, 2002-435328/46.
XX
XX
XX Detecting granulocyte activation by detecting differential expression of
XX genes associated with granulocyte activation, which serves as diagnostic
XX markers that is useful for monitoring disease states and drug toxicity.
XX
XX
XX Claim 1; SEQ ID NO 279; 114pp; English.
XX
XX The invention relates to detecting (M1) granulocyte (GC) activation
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
XX DNA chip analysis as given in the specification, and comparing the
XX expression level to an expression level in an unactivated GC, where
XX differential expression of Gs is indicative of GCA. Also included are
XX modulating (M2) GA by contacting GC with an agent that alters the
XX expression of at least one gene in Gs; (2) screening (M3) for an agent
XX capable of modulating GCA or an inflammation (especially chronic) in a
XX tissue, an allergic response in a subject, exposure of a subject to a
XX pathogen or sterile inflammatory disease using the gene expression
XX profile; (3) detecting (M4) an inflammation (especially chronic) in a
XX tissue, an allergic response in a subject, exposure of a subject to a
XX pathogen or sterile inflammatory disease, by detecting the level of
XX expression in a sample of the tissue of gene(s) from Gs, where the level
XX of expression of the gene is indicative of inflammation; (4) treating
XX (M5) an inflammation (especially chronic) or in a tissue, an allergic
XX response in a subject, exposure of a subject to a pathogen or sterile
XX inflammatory disease, by contacting a tissue having inflammation with an
XX agent that modulates the expression of gene(s) from Gs in the tissue. M1
XX is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
XX for screening an agent capable of modulating GCA preferably in an
XX inflammation in a tissue; M4 is useful for detecting an inflammation
XX (especially chronic) in a tissue, an allergic response in a subject,
XX exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
XX psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
XX cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
XX respiratory distress syndrome, inflammatory bowel disease, Crohn's
XX disease, ulcerative colitis, periodontal disease; also bacterial
XX infection, viral infection, parasitic infection, protozoal infection,
XX fungal infection and M5 is useful for treating one of the above
XX conditions. The present sequence represents a gene differentially
XX expressed in granulocytes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 6571 BP; 2335 A; 1164 C; 1415 G; 1657 T; 0 U; 0 Other;

Alignment Scores:
Score: 9.14e+03 Length: 6571
Prod. No.: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x ABK83708 (1-6571)

QY 1 GlyTyr**ValGluGlu 6
DB 3061 GGATATTCTGTGAGAGAA 3078

Search completed: November 3, 2004, 15:20:13
Job time : 361 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2004, 13:32:27 ; Search time 2509 Seconds

(without alignments)
87.142 Million cell updates/sec

Title: US-10-030-194A-5

Perfect score: 28

Sequence: 1 GYXVEE 6

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10030194/runat_01112004.184847.28965/app.query.fasta_1.398
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=1000
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10030194@cgn 1.1.7873@runat_01112004.184847.28965 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	27	96.4	118	5	BU016086
2	27	96.4	131	5	BQ976013
3	27	96.4	131	5	BQ978589
4	27	96.4	132	5	BU020428
5	27	96.4	138	5	BQ912701
6	27	96.4	151	2	BF047816
7	27	96.4	161	6	CD729971
C 8	27	96.4	162	8	BH316542
9	27	96.4	175	5	BU017076

10	27	96.4	175	6	CD972375
C 11	27	96.4	175	6	CD972866
C 12	27	96.4	175	6	CD976726
C 13	27	96.4	179	9	CG512665
C 14	27	96.4	185	4	B1177133
C 15	27	96.4	185	4	B1177140
C 16	27	96.4	195	6	CA904403
C 17	27	96.4	201	1	AA550752
18	27	96.4	204	5	BQ912607
19	27	96.4	217	2	BBS72901
C 20	27	96.4	217	6	CB129746
C 21	27	96.4	218	5	BU023155
C 22	27	96.4	221	5	BF509543
C 23	27	96.4	228	5	BU020431
C 24	27	96.4	231	8	BZ090362
C 25	27	96.4	232	5	BU018501
C 26	27	96.4	239	5	BQ968398
C 27	27	96.4	244	1	AA301321
C 28	27	96.4	247	2	AA477834
C 29	27	96.4	261	8	BZ260573
C 30	27	96.4	261	9	CG505068
C 31	27	96.4	262	7	CR521180
C 32	27	96.4	264	1	AA334794
C 33	27	96.4	270	9	AG228772
C 34	27	96.4	271	1	AV278603
C 35	27	96.4	277	2	BB557903
C 36	27	96.4	279	2	BB516647
C 37	27	96.4	281	5	BQ911194
C 38	27	96.4	284	9	CG502243
C 39	27	96.4	287	5	BU019672
C 40	27	96.4	287	6	CF034190
C 41	27	96.4	287	9	CG502254
C 42	27	96.4	289	1	AV144507
C 43	27	96.4	293	4	B1046607
C 44	27	96.4	295	7	CK248536
C 45	27	96.4	298	2	AW086727
C 46	27	96.4	298	6	CA516999
C 47	27	96.4	300	5	BU017825
C 48	27	96.4	303	2	AW098038
C 49	27	96.4	303	2	BB531239
C 50	27	96.4	303	6	CB700488
C 51	27	96.4	304	7	CO415454
C 52	27	96.4	308	7	CR469122
C 53	27	96.4	309	8	AQ233114
C 54	27	96.4	317	1	AI715497
C 55	27	96.4	317	6	CF034326
C 56	27	96.4	320	5	BQ146842
C 57	27	96.4	321	5	BQ911150
C 58	27	96.4	322	6	CD736806
C 59	27	96.4	324	5	BQ910364
C 60	27	96.4	333	6	CD476039
C 61	27	96.4	333	7	R31872
C 62	27	96.4	335	5	BQ974743
C 63	27	96.4	337	5	BU016795
C 64	27	96.4	338	4	BF989730
C 65	27	96.4	339	4	BF989729
C 66	27	96.4	339	5	BU017766
C 67	27	96.4	341	5	BQ975877
C 68	27	96.4	349	5	BU024708
C 69	27	96.4	350	5	BU025543
C 70	27	96.4	351	5	BQ616698
C 71	27	96.4	351	5	BU018981
C 72	27	96.4	353	2	AW294420
C 73	27	96.4	355	5	BQ910337
C 74	27	96.4	356	5	BF993748
C 75	27	96.4	356	5	BU018118
C 76	27	96.4	356	9	CR520931
C 77	27	96.4	356	9	CG644714
C 78	27	96.4	358	1	AI509762
C 79	27	96.4	358	5	BQ975182
C 80	27	96.4	364	5	BQ978782
C 81	27	96.4	366	5	BQ977034
C 82	27	96.4	367	1	AL044871

CD972375	QAE20h02.
CD972866	QAE27g02.
CD976726	QAF21a03.
CG512665	OST65684
B1177133	EST518078
B1177140	EST518085
CA904403	FCSC13986
AA550752	rj84f09.8
BQ912607	QBA21H01.
BBS72901	BB572901
CB129746	K-EST0179
BU023155	QHE9L24.Y
BF509543	UI-H-B14-
BU020431	QHE27G17.
BZ090362	CH230-214
BU018501	QHE18N13.
BQ968398	QHB33M10.
AA301321	EST14252
AA477834	17053 MAR
BZ260573	CH230-285
CG505068	OST53833
CR521180	CR521180
AA334794	EST33059
AG228772	Lotus cor
AV278603	AV278603
BB557903	BB557903
BB516647	BB516647
BQ911194	QHA16G21.
CG502243	OST46996
BU019672	QHE22M17.
CF034190	QCF2c02.Y
CG502254	OST47010
AV144507	AV144507
B1046607	MR3-FN020
CK248536	EST732173
AW086727	ga04e02.X
CA516999	KS09068B1
BU017825	QHE16P11.
AW098038	ga04e02.Y
BB531239	BB531239
CB700488	AMGNNUC.N
CO415454	Mdf33006d
CR469122	CR469122
AQ233114	HS_2037.A
AI715497	UI-R-YO-a
CF034326	QCF3h10.Y
BQ146842	NF029A03F
BQ911150	QHA16D20.
CD736806	4020688.1
BQ910364	QHA13P08.
CD476039	rad03-2ms
R31872	Yh59e10.r1
BQ974743	QHI16124
BU016795	QHE14A15.
BF989730	IL5-GN017
BF989729	IL5-GN017
BU017766	QHE16M20.
BQ975877	QHI1M14.Y
BU024708	QHFSL21.Y
BU025543	QHFSP05.Y
BQ616698	faa94a09.
BU018981	QHE1907.Y
AW294420	UI-H-B12-
BQ910337	QHA13002.
BF993748	IL5-GN017
BU018118	QHE17L22.
CR520931	CR520931
CG644714	OST388385
AI509762	vx24e07.Y
BQ975182	QHI18K19.
BQ978782	QHI5E02.Y
BQ977034	QHI3L08.
AL044871	DKF2p434F

83	27	96.4	367	5	BQ916909	BQ916909 QHB19006.	156	27	96.4	445	5	BQ969599	BQ969599 QHB39A04.
84	27	96.4	367	5	BQ928527	BQ928527 QHB15P01.	157	27	96.4	445	8	BH399653	BH399653 AG-ND-119
85	27	96.4	368	5	BQ979888	BQ979888 QH15J01.Y	158	27	96.4	446	1	AA867196	AA867196 vx24e07.r
86	27	96.4	369	5	BQ923893	BQ923893 QHB14J12.	159	27	96.4	446	5	BQ913048	BQ913048 QH6A02.Y
87	27	96.4	369	6	CD976494	CD976494 QAF18B12.	160	27	96.4	446	5	BQ970858	BQ970858 QHB8A23.Y
88	27	96.4	371	5	BQ975786	BQ975786 QH1105.Y	161	27	96.4	447	5	BQ912659	BQ912659 QHA12P11.
89	27	96.4	372	5	BQ252801	BQ252801 san81a11.	162	27	96.4	447	5	BQ914096	BQ914096 QHA9F03.Y
90	27	96.4	372	5	BQ334049	BQ334049 QHJ2P07.Y	163	27	96.4	447	5	BQ916289	BQ916289 QHE12K18.
91	27	96.4	372	8	AO651125	AO651125 Sheared.D	164	27	96.4	447	7	CK613436	CK613436 LPSK N04
92	27	96.4	374	4	EM060338	EM060338 KS01013C0	165	27	96.4	451	5	BQ966522	BQ966522 QHB27B19.
93	27	96.4	374	5	BQ978669	BQ978669 QH15N11.Y	166	27	96.4	451	5	BQ916663	BQ916663 QHE13K22.
94	27	96.4	375	1	AA976900	AA976900 cq35e01.s	167	27	96.4	451	7	CN579531	CN579531 Mdfw2033f
95	27	96.4	376	1	A1653372	A1653372 wb45d01.x	168	27	96.4	452	7	N94038	N94038 za24e06.r1
96	27	96.4	376	1	A1989917	A1989917 wt49c10.x	169	27	96.4	454	1	AI164099	AI164099 A054p16u
97	27	96.4	376	1	CD411942	CD411942 Gm ck4293	170	27	96.4	454	4	BG131385	BG131385 EST46277
98	27	96.4	383	9	CG554314	CG554314 OST167779	171	27	96.4	456	6	CB740445	CB740445 AMGNNUC.N
99	27	96.4	384	8	BH858437	BH858437 B5_40482	172	27	96.4	457	7	CF422643	CF422643 NCST3378
100	27	96.4	385	5	BY153719	BY153719 BY153719-	173	27	96.4	458	5	BQ916399	BQ916399 QHB17M05.
101	27	96.4	386	1	AA014084	AA014084 mh28c03.r	174	27	96.4	458	5	BQ916399	BQ916399 QHB17M05.
102	27	96.4	387	1	A1103564	A1103564 EST212853	175	27	96.4	462	9	AG268135	AG268135 Cyanid108
103	27	96.4	387	9	CL444076	CL444076 ZMBB046	176	27	96.4	463	8	A2142018	A2142018 SP_0008.A
104	27	96.4	389	5	BQ856597	BQ856597 QG5E05.Y	177	27	96.4	467	1	AI015194	AI015194 ct82g04.r6
105	27	96.4	390	2	BE017313	BE017313 fk81c01.Y	178	27	96.4	467	5	BQ045607	BQ045607 EST594725
106	27	96.4	391	9	BQ967648	BQ967648 OST461640	179	27	96.4	468	5	BQ916399	BQ916399 QHB17M05.
107	27	96.4	392	5	BQ975613	BQ975613 QH119P24.	180	27	96.4	468	7	CK433208	CK433208 QH0014B.T
108	27	96.4	392	5	BQ975613	BQ975613 QH119P24.	181	27	96.4	469	5	BQ969740	BQ969740 QHB39G15.
109	27	96.4	392	8	AX921291	AX921291 BX921291	182	27	96.4	469	5	BQ977048	BQ977048 QH123L24.
110	27	96.4	392	8	AX921291	AX921291 RPCT-23-2	183	27	96.4	472	5	BQ914084	BQ914084 QH49E13.Y
111	27	96.4	393	2	CO416625	CO416625 Mdf3008C	184	27	96.4	475	4	BG144478	BG144478 ut65e11.x
112	27	96.4	394	2	BE301622	BE301622 bf75g07.x	185	27	96.4	475	5	BQ910184	BQ910184 QH413H08.
113	27	96.4	394	5	BX322450	BX322450 PIC1_5.A0	186	27	96.4	475	6	CA569147	CA569147 K0438E01-
114	27	96.4	395	5	BX669832	BX669832 BX669832	187	27	96.4	476	5	BQ910779	BQ910779 QH415C32.
115	27	96.4	398	8	A2811084	A2811084 2M0077L03	188	27	96.4	476	8	AO940743	AO940743 Sheared.D
116	27	96.4	400	6	CF089775	CF089775 QH23P09.	189	27	96.4	478	6	CD319399	CD319399 StrPu538.
117	27	96.4	403	4	BF992503	BF992503 ILS-GN017	190	27	96.4	478	6	AZ031627	AZ031627 RPCT-23-2
118	27	96.4	404	8	AQ970039	AQ970039 RPCT-23-3	191	27	96.4	479	5	BE467914	BE467914 hz75g05.x
119	27	96.4	407	4	BF993065	BF993065 ILS-GN017	192	27	96.4	479	5	BQ913988	BQ913988 QH49A08.Y
120	27	96.4	408	6	CB927385	CB927385 ABAL_25.C	193	27	96.4	479	1	AI892143	AI892143 mh28c03.Y
121	27	96.4	412	2	AN570472	AN570472 sj24a01.Y	194	27	96.4	480	5	BQ9667450	BQ9667450 A0853F11-
122	27	96.4	417	2	CN238041	CN238041 RJAL15E12	195	27	96.4	480	5	BQ919163	BQ919163 QHE20H01.
123	27	96.4	421	1	A1212520	A1212520 x7h06a1.r	196	27	96.4	481	9	CG663862	CG663862 OST449835
124	27	96.4	421	2	AW522103	AW522103 UI-R-BJ0p	197	27	96.4	484	9	CG978743	CG978743 CH240.L70
125	27	96.4	421	7	CO513659	CO513659 sl3dSG17D	198	27	96.4	486	5	BQ910828	BQ910828 QH415E24.
126	27	96.4	424	1	A1023281	A1023281 ov65h07.s	199	27	96.4	489	1	AV623097	AV623097 QHA15E24.
127	27	96.4	426	5	BQ419403	BQ419403 faa38f12.	200	27	96.4	489	5	BQ910828	BQ910828 QH415E24.
128	27	96.4	427	5	BQ915735	BQ915735 QHB15J21.	201	27	96.4	490	1	AA013513	AA013513 mh25e06.r
129	27	96.4	427	7	CN237912	CN237912 WLB142D04.	202	27	96.4	490	4	BQ980074	BQ980074 BJ498074
130	27	96.4	430	8	B55124	B55124 CTT-HSP-386	203	27	96.4	491	7	CK697855	CK697855 ZF101-F00
131	27	96.4	431	4	BG149132	BG149132 74_L1N01	204	27	96.4	492	1	AI552551	AI552551 vx24e07.x
132	27	96.4	431	5	BQ021776	BQ021776 QHE4e02.Y	205	27	96.4	492	1	BE115974	BE115974 UI-R-B81-
133	27	96.4	431	6	C95283	C95283 C95283_C1r	206	27	96.4	493	1	AV994978	AV994978 AV994978
134	27	96.4	434	1	AA885084	AA885084 am30h01.s	207	27	96.4	493	2	BE110120	BE110120 UI-R-BJ1-
135	27	96.4	434	9	CG545897	CG545897 OST144662	208	27	96.4	496	2	BE600528	BE600528 P11_94_C1
136	27	96.4	435	5	BQ020310	BQ020310 QHE27A14.	209	27	96.4	497	2	BE646993	BE646993 UI-M-BH1-
137	27	96.4	435	5	BQ022886	BQ022886 QHB8P10.Y	210	27	96.4	498	1	AA964455	AA964455 UI-R-E1-g
138	27	96.4	436	5	BQ913507	BQ913507 QHA7K18.Y	211	27	96.4	498	1	AL368637	AL368637 mtBA25G02
139	27	96.4	437	1	AJ541701	AJ541701 AJ541701	212	27	96.4	499	2	BF706129	BF706129 280377.MA
140	27	96.4	437	5	BQ911521	BQ911521 QH417G23.	213	27	96.4	499	8	AO515819	AO515819 HS_5237.A
141	27	96.4	438	5	BQ661887	BQ661887 HP01L05W	214	27	96.4	500	5	BQ912116	BQ912116 QH41d06.Y
142	27	96.4	438	5	BQ017794	BQ017794 QHE16N24.	215	27	96.4	501	6	CA566361	CA566361 Q0400D04-
143	27	96.4	438	5	BQ028738	BQ028738 QH33h03.Y	216	27	96.4	502	7	CK443206	CK443206 KQ0014B.B
144	27	96.4	439	1	AI178505	AI178505 mt14g07.r	217	27	96.4	503	1	AV977694	AV977694 AV977694
145	27	96.4	440	4	BI189574	BI189574 f3b07fs.r	218	27	96.4	504	9	CG616571	CG616571 OST308913
146	27	96.4	442	7	CN956564	CN956564 4064_5001	219	27	96.4	505	9	CG595692	CG595692 OST255867
147	27	96.4	443	5	BQ911046	BQ911046 QH415011.	220	27	96.4	506	2	BE100310	BE100310 UI-R-BJ1-
148	27	96.4	443	5	BQ913679	BQ913679 QH48C10.Y	221	27	96.4	507	6	CF122959	CF122959 UI-HF-CB0
149	27	96.4	443	5	BQ431758	BQ431758 UI-HF-ENO	222	27	96.4	508	9	CR050610	CR050610 Reverse.s
150	27	96.4	443	6	CD730876	CD730876 4039796_1	223	27	96.4	508	9	PT015A23R	PT015A23R Parametru
151	27	96.4	443	8	B74760	B74760 CTT-HSP-204	224	27	96.4	510	7	CN514592	CN514592 2444.1.Tu
152	27	96.4	443	8	BZ497185	BZ497185 BOND50TR	225	27	96.4	511	5	BQ121126	BQ121126 EST606702
153	27	96.4	444	2	B8814189	B8814189 BB814189	226	27	96.4	511	5	BQ190007	BQ190007 UI-R-CN1-
154	27	96.4	444	5	BQ911838	BQ911838 QH419E12.	227	27	96.4	511	5	BQ749756	BQ749756 SNEST4a72
155	27	96.4	444	5	BQ965400	BQ965400 QHB21L04.	228	27	96.4	511	7	CO778181	CO778181 BL003A_F1

375	27	96.4	625	4	BG710506	pglin.pk0	448	27	96.4	665	6	CD848246	CD848246	DH0AC003Z
376	27	96.4	626	6	CB497344	omykrbna5	C 449	27	96.4	665	8	A2360567	A2360567	LM0103B23
377	27	96.4	626	6	CB687027	CEST-07-E	C 450	27	96.4	667	2	B2467004	UI-N-CQOP	UI-N-CQOP
378	27	96.4	626	7	CF443077	EST679422	451	27	96.4	667	5	BW277293	BW277293	BW277293
379	27	96.4	626	8	BZ676216	PUBET43TD	452	27	96.4	667	7	CN883758	CN883758	010814AAS
380	27	96.4	627	2	BE801355	sr13e09.Y	C 453	27	96.4	667	9	CR245155	Forward s	CR245155
381	27	96.4	627	5	BW317649	BW317649	C 454	27	96.4	668	8	AQ447486	mgx50007H	AQ447486
382	27	96.4	628	4	B1203362	EST521402	C 455	27	96.4	669	1	AV871671	AV871671	AV871671
383	27	96.4	628	4	BJ010824	BJ010824	C 456	27	96.4	672	2	AW219103	AW219103	AW219103
384	27	96.4	629	5	BF146304	BF146304	C 457	27	96.4	672	6	CB423686	CB423686	CB423686
385	27	96.4	629	5	CG912136	ZMBB036	C 458	27	96.4	672	7	CN544128	UI-R-D21	UI-R-D21
386	27	96.4	630	4	BJ002074	BJ002074	C 459	27	96.4	673	4	BN859185	fy57g12.x	fy57g12.x
387	27	96.4	630	4	BW491002	p2p2n.pk0	C 460	27	96.4	675	4	BG168663	602339682	602339682
388	27	96.4	630	5	BU048680	PP_LEA003	C 461	27	96.4	675	6	CD441772	EL01N0563	EL01N0563
389	27	96.4	630	6	BY722109	BY722109	C 462	27	96.4	677	6	CD763012	GGEZSM103	GGEZSM103
390	27	96.4	631	4	B1205594	EST523634	C 463	27	96.4	677	9	AG132383	Pan_crog1	Pan_crog1
391	27	96.4	631	4	B1206586	EST524626	464	27	96.4	678	5	BU489664	604131752	604131752
392	27	96.4	631	5	BU286568	604163133	465	27	96.4	678	5	BU490270	604131414	604131414
393	27	96.4	631	9	CC991409	ZUAEPS8TH	C 466	27	96.4	678	6	CA904405	PCSC18062	PCSC18062
394	27	96.4	632	1	AT600085	EST251788	C 467	27	96.4	679	5	BW219263	BW219263	BW219263
395	27	96.4	633	1	AI161566	A002P61U	C 468	27	96.4	680	4	BJ726495	BJ726495	BJ726495
396	27	96.4	635	4	BI994449	1031008G0	469	27	96.4	681	4	BM427553	pgfzn.pk0	pgfzn.pk0
397	27	96.4	635	9	CE736265	tigr-g8s-	C 470	27	96.4	681	5	BQ403677	GA_Ed006	GA_Ed006
398	27	96.4	636	7	CN914235	021210ABN	C 471	27	96.4	681	7	CN848350	PG07023D0	PG07023D0
399	27	96.4	636	7	CO028303	EST806687	C 472	27	96.4	682	4	BI296303	UI-R-DK0-	UI-R-DK0-
400	27	96.4	637	2	BE608348	sq29b08.Y	C 473	27	96.4	682	5	BQ004395	UI-H-E10-	UI-H-E10-
401	27	96.4	637	5	BN979724	UI-CF-DU1	C 474	27	96.4	683	5	BU137778	603122767	603122767
402	27	96.4	639	5	BU486873	604125375	475	27	96.4	683	5	BU360311	603407693	603407693
403	27	96.4	640	8	BH065685	RPCI-24-2	C 476	27	96.4	683	6	CB165964	RKT603019	RKT603019
404	27	96.4	640	9	CC790592	ZMBBB016	C 477	27	96.4	685	1	AV886757	AV886757	AV886757
405	27	96.4	641	1	AJ805100	AJ805100	C 478	27	96.4	685	4	BM519372	07-C1-RK2E	07-C1-RK2E
406	27	96.4	641	4	BJ701775	BJ701775	C 479	27	96.4	685	6	CD821573	BN25_042E	BN25_042E
407	27	96.4	642	4	BU565176	BU565176	C 480	27	96.4	686	5	BU258825	603413037	603413037
408	27	96.4	642	6	CA360947	634679.NC	481	27	96.4	686	5	BU287693	604166368	604166368
409	27	96.4	643	2	AW132771	se10a06.Y	C 482	27	96.4	687	5	BU491685	604131844	604131844
410	27	96.4	643	4	BJ019971	BJ019971	C 483	27	96.4	688	7	CN293251	BW312981	BW312981
411	27	96.4	645	4	BJ612240	BJ612240	484	27	96.4	688	7	CN293251	170006002	170006002
412	27	96.4	645	5	BU884774	BU884774	C 485	27	96.4	689	4	BJ516685	BJ516685	BJ516685
413	27	96.4	645	7	CN984039	BN844774	C 486	27	96.4	689	7	CN237348	WLB133807	WLB133807
414	27	96.4	646	1	AJ802386	AJ802386	C 487	27	96.4	690	8	BH117682	RPCI-24-3	RPCI-24-3
415	27	96.4	646	6	CD575912	UCRPT01_0	C 488	27	96.4	691	5	BU884159	R006G03_P	R006G03_P
416	27	96.4	647	5	BU480721	603847208	C 489	27	96.4	691	5	BW081952	BW081952	BW081952
417	27	96.4	647	8	AZ572651	308PVD10	C 490	27	96.4	691	6	BY710126	BY710126	BY710126
418	27	96.4	648	1	AA660732	00623_McR	C 491	27	96.4	691	6	CB502411	sealplnb5	sealplnb5
419	27	96.4	648	5	BF461619	BP461619	C 492	27	96.4	693	4	BU016605	BU016605	BU016605
420	27	96.4	648	6	CA359801	632871.NC	C 493	27	96.4	693	7	CR435042	CR435042	CR435042
421	27	96.4	648	6	CB940896	IPCGJx14	C 494	27	96.4	694	5	BU479684	603842493	603842493
422	27	96.4	649	5	BP455201	BP455201	C 495	27	96.4	694	5	BW216142	BW216142	BW216142
423	27	96.4	649	5	BU486465	BU486465	496	27	96.4	694	5	BX252073	BX252073	BX252073
424	27	96.4	650	1	AV961458	AV961458	497	27	96.4	694	9	CR210170	Reverse s	Reverse s
425	27	96.4	650	4	BM860883	fy51f10.Y	C 498	27	96.4	695	4	BJ522161	BJ522161	BJ522161
426	27	96.4	651	9	CL193343	104_416_1	C 499	27	96.4	695	7	CF768025	CES005202	CES005202
427	27	96.4	652	6	CA118781	SCBGLR11	C 500	27	96.4	695	9	CL696966	SP_Ba000	SP_Ba000
428	27	96.4	653	4	BI270952	NF056F07F	C 501	27	96.4	696	5	BU331583	603501665	603501665
429	27	96.4	653	6	BY722278	BY722278	502	27	96.4	696	5	BU884826	R016C04_P	R016C04_P
430	27	96.4	654	7	CF418004	USDA-FP_1	C 503	27	96.4	696	7	CF418289	USDA-FP_1	USDA-FP_1
431	27	96.4	654	7	BW213076	BW213076	C 504	27	96.4	697	5	BU491100	604130344	604130344
432	27	96.4	656	8	BH466492	BOGULS9TR	C 505	27	96.4	697	9	AG138176	Pan_crog1	Pan_crog1
433	27	96.4	659	2	BB627757	BB627757	C 506	27	96.4	698	5	BU869928	Q006C02_P	Q006C02_P
434	27	96.4	659	5	BU072748	BU072748	C 507	27	96.4	698	7	CO079842	GR_Ea420	GR_Ea420
435	27	96.4	660	4	BJ001321	BJ001321	C 508	27	96.4	699	7	AG344797	USF_muscu	USF_muscu
436	27	96.4	660	4	BJ535713	BJ535713	C 509	27	96.4	699	9	CF418221	CF418221	CF418221
437	27	96.4	660	5	BQ415065	BQ415065	C 510	27	96.4	700	1	AV866561	AV866561	AV866561
438	27	96.4	660	5	BX850154	BX850154	511	27	96.4	701	1	AV866561	AV866561	AV866561
439	27	96.4	661	4	BJ580425	BJ580425	C 512	27	96.4	702	5	BW147680	BW147680	BW147680
440	27	96.4	661	6	CB510543	sealnhw50	513	27	96.4	704	5	BU487860	604127209	604127209
441	27	96.4	662	7	CO867992	Mdfct3037	514	27	96.4	705	5	BU004979	QGG6M02.Y	QGG6M02.Y
442	27	96.4	663	4	BJ513686	BJ513686	515	27	96.4	705	9	BU223298	603797587	603797587
443	27	96.4	663	5	BQ414910	GA_Ed009	516	27	96.4	705	9	CL208143	ZMBB0055	ZMBB0055
444	27	96.4	663	9	CC766091	CH240_131	C 517	27	96.4	706	5	BU384454	BU384454	BU384454
445	27	96.4	664	4	BI721952	1031059A1	C 518	27	96.4	707	5	BW217732	BW217732	BW217732
446	27	96.4	664	7	CN526032	UI-N-HN0-	C 519	27	96.4	708	2	BB635737	BB635737	BB635737
447	27	96.4	664	8	BZ643121	OGCAF69TM	C 520	27	96.4	708	4	BM075042	MEST349-D	MEST349-D

521	27	96.4	708	6	CA145273	SCSRT206	594	27	96.4	765	2	BF861441	BF861441	963023B01
522	27	96.4	708	7	CN293252	170004245	595	27	96.4	765	5	BU107416	BU107416	603111591
523	27	96.4	709	8	AZ735612	RPCC1-24-1	596	27	96.4	765	5	BU490045	BU490045	604132102
524	27	96.4	709	8	BH688339	BMJN785TR	597	27	96.4	766	5	EX858533	EX858533	EX858533
525	27	96.4	710	5	EX956231	DKP2p781C	598	27	96.4	767	4	CG910481	CG910481	602806157
526	27	96.4	710	9	CR147362	Reverse B	c 599	27	96.4	768	1	AU000335	AU000335	AU000335
527	27	96.4	713	6	CD852334	DH0ALL272	c 600	27	96.4	768	9	CG809058	CG809058	ESAB181F
528	27	96.4	716	7	CK026884	AGENCOURT	c 601	27	96.4	769	4	BG647013	BG647013	EST508632
529	27	96.4	717	6	CD848287	DH0AC004Z	c 602	27	96.4	769	6	CA386667	CA386667	668311 NC
530	27	96.4	718	4	BI907792	603066043	603	27	96.4	770	5	BU488724	BU488724	604126158
531	27	96.4	718	4	BU724999	BJ724999	604	27	96.4	771	7	CN510376	CN510376	AGENCOURT
532	27	96.4	718	5	BU490302	604129414	c 605	27	96.4	771	9	CR184418	CR184418	Reverse B
533	27	96.4	718	5	EX298136	EX298136	606	27	96.4	771	9	CG386948	CG386948	ZMMBBC056
534	27	96.4	718	6	CF178946	813024 MA	607	27	96.4	772	5	EX851281	EX851281	EX851281
535	27	96.4	718	6	AG426620	Mus muscu	608	27	96.4	774	9	CL695961	CL695961	PR1017 B
536	27	96.4	719	9	CL477583	SAIL 275	609	27	96.4	776	6	CA918139	CA918139	EST642286
537	27	96.4	720	4	BI733473	603352525	610	27	96.4	777	7	CF999437	CF999437	AGENCOURT
538	27	96.4	721	9	CR062758	Forward B	611	27	96.4	778	7	CN037807	CN037807	nm 24 c5
539	27	96.4	722	4	BI310411	EST531216	c 612	27	96.4	779	9	CR094829	CR094829	Reverse B
540	27	96.4	722	6	CD760806	GGEZSM100	613	27	96.4	780	4	BG199309	BG199309	RST18591
541	27	96.4	722	7	CN152892	939933 MA	614	27	96.4	781	7	CN520475	CN520475	Q00107 B3
542	27	96.4	722	7	CN154972	942237 MA	615	27	96.4	782	5	EX298986	EX298986	EX298986
543	27	96.4	722	8	AZ309611	1M0016D02	616	27	96.4	783	4	BJ715170	BJ715170	BJ715170
544	27	96.4	723	4	BU726513	BU726513	617	27	96.4	783	7	CK695037	CK695037	ZF101-P00
545	27	96.4	723	9	AG380204	Mus muscu	618	27	96.4	783	7	CO810486	CO810486	AGENCOURT
546	27	96.4	724	2	BF209055	601873149	c 619	27	96.4	784	5	EX072750	EX072750	603065236
547	27	96.4	724	4	BU710897	BU710897	620	27	96.4	784	5	BI911804	BI911804	BX072750
548	27	96.4	725	9	AG419747	Mus muscu	621	27	96.4	784	6	CD486211	CD486211	CFU57.3B1
549	27	96.4	726	4	BU722161	BU722161	622	27	96.4	784	8	B2168746	B2168746	CD230-267
550	27	96.4	726	6	CD761433	GGEZSM101	623	27	96.4	785	6	CB906845	CB906845	tri-c080x1
551	27	96.4	727	2	BF219565	GM700018A	624	27	96.4	786	9	BG167318	BG167318	Danilo rcr
552	27	96.4	727	4	BU736601	BU736601	625	27	96.4	787	4	BG936510	BG936510	Ssl-0885
553	27	96.4	727	4	BM680082	UI-E-E01-	c 626	27	96.4	787	6	CF257738	CF257738	AGENCOURT
554	27	96.4	728	9	CG297117	OG5CX88TC	627	27	96.4	787	6	CF257738	CF257738	phao14 h0
555	27	96.4	729	5	BP694481	BP694481	628	27	96.4	788	5	BW089292	BW089292	BW089292
556	27	96.4	732	5	BU326161	603491826	629	27	96.4	788	9	EX218802	EX218802	Danilo rcr
557	27	96.4	732	7	CF879397	tr1c080x1	c 630	27	96.4	789	4	BU814510	BU814510	BU814510
558	27	96.4	736	4	BI819150	603036948	c 631	27	96.4	789	7	CK178003	CK178003	EST767323
559	27	96.4	737	4	BM017631	603644844	632	27	96.4	789	9	CG06589	CG06589	OGW1377V
560	27	96.4	737	5	BU026533	BU026533	c 633	27	96.4	790	4	BG622496	BG622496	602647222
561	27	96.4	738	4	CG590959	EST498801	634	27	96.4	792	5	BU341626	BU341626	603519964
562	27	96.4	738	9	AG426192	AG426192	c 635	27	96.4	792	7	CK778724	CK778724	965893 MA
563	27	96.4	739	1	AJ454420	AJ454420	c 636	27	96.4	792	7	CO112171	CO112171	GR_Eb004
564	27	96.4	739	4	BU714282	BU714282	c 637	27	96.4	792	8	B2071620	B2071620	1julic09.
565	27	96.4	740	4	BI714300	BI714300	638	27	96.4	794	5	BU477477	BU477477	603841690
566	27	96.4	741	4	BI203201	EST521241	639	27	96.4	794	5	BU488994	BU488994	604125522
567	27	96.4	741	5	BU400126	604141088	640	27	96.4	794	8	B2524250	B2524250	OGAJA567C
568	27	96.4	741	8	CC111448	NDL13J3.	641	27	96.4	796	4	BG476264	BG476264	602525129
569	27	96.4	742	2	BE822458	BE822458	c 642	27	96.4	796	7	CO109259	CO109259	GR_Eb004
570	27	96.4	742	5	EX317585	EX317585	c 643	27	96.4	797	6	CF204312	CF204312	RR8909151
571	27	96.4	742	6	CD851551	DH0ALL14Z	644	27	96.4	797	9	CG812479	CG812479	ZMMBBC052
572	27	96.4	742	7	CF996096	CF996096	645	27	96.4	799	8	B2428063	B2428063	BONSJ84TF
573	27	96.4	743	9	AG482895	AG482895	c 646	27	96.4	799	9	CG674596	CG674596	tr82328 t
574	27	96.4	743	9	CE623890	tigr-g88-	647	27	96.4	800	7	CN836099	CN836099	AGENCOURT
575	27	96.4	744	6	CA347743	678913 NC	c 648	27	96.4	803	7	CF443847	CF443847	EST680192
576	27	96.4	746	9	CG729565	OGUDW64TV	c 649	27	96.4	803	9	EX982635	EX982635	Reverse B
577	27	96.4	747	9	CG912157	t070g02ba	c 650	27	96.4	803	9	CG123893	CG123893	Reverse B
578	27	96.4	749	5	BU251670	BU251670	c 651	27	96.4	803	9	CG109824	CG109824	PUI0U187B
579	27	96.4	750	7	CN013465	AGENCOURT	c 652	27	96.4	804	9	EX963963	EX963963	Reverse B
580	27	96.4	751	5	BU220805	BU220805	c 653	27	96.4	805	8	CO061427	CO061427	MUGO_CH25
581	27	96.4	753	6	BY736531	BY736531	c 654	27	96.4	807	5	BU250421	BU250421	603402042
582	27	96.4	756	7	CF549634	AGENCOURT	c 655	27	96.4	810	2	BF971375	BF971375	602273059
583	27	96.4	758	2	BE614526	BE614526	c 656	27	96.4	810	8	BZ703960	BZ703960	PUCNC13TD
584	27	96.4	758	4	BI762839	603048371	c 657	27	96.4	810	9	CF090040	CF090040	Reverse B
585	27	96.4	759	5	BU830244	T005G06 F	c 658	27	96.4	811	9	CG225526	CG225526	OG2AM68TH
586	27	96.4	759	5	AG583584	Mus muscu	c 659	27	96.4	813	7	CO070898	CO070898	GR_Ea2BD
587	27	96.4	760	7	CO572343	AGENCOURT	c 660	27	96.4	814	7	CK635021	CK635021	UI-M-HN0-
588	27	96.4	761	5	BU115105	BU115105	c 661	27	96.4	817	8	BZ757695	BZ757695	PUF8H37TD
589	27	96.4	761	8	BZ062684	lle42a11.	662	27	96.4	818	7	CK029661	CK029661	AGENCOURT
590	27	96.4	762	9	CL672710	PR1017C_E	663	27	96.4	820	7	CK362078	CK362078	AGENCOURT
591	27	96.4	763	5	EX453506	EX453506	c 664	27	96.4	824	6	CD575844	CD575844	UCRPT01_0
592	27	96.4	764	5	EX321096	EX321096	c 665	27	96.4	827	7	CF548978	CF548978	AGENCOURT
593	27	96.4	764	5	EX321096	EX321096	666	27	96.4					

667	27	96.4	827	8	B2720272	PUBDA78TD	740	27	96.4	913	8	B2754596	PUBEH65TD
668	27	96.4	830	9	CC642772	OGUGV03TV	741	27	96.4	913	9	CC619302	AGUFU38TH
669	27	96.4	831	9	BU694378	BU694378	742	27	96.4	914	7	CN501429	AGENCOURT
670	27	96.4	833	8	B2702120	PUCN37TD	743	27	96.4	914	7	CO773189	testis_ES
671	27	96.4	834	7	CG398643	602440248	744	27	96.4	915	6	CD327801	AGENCOURT
672	27	96.4	834	7	CO119827	GR_EB022	745	27	96.4	916	5	BU109754	603125524
673	27	96.4	835	5	BU114403	603130234	746	27	96.4	917	7	CO809863	AGENCOURT
674	27	96.4	835	5	CG813786	ZMBBB051	747	27	96.4	922	2	BF346822	602021616
675	27	96.4	838	5	BU492390	604132379	748	27	96.4	923	7	CN972808	20305_124
676	27	96.4	838	5	CD574198	UCRPT01_0	749	27	96.4	930	5	CNS021Y3	Tetraodon
677	27	96.4	840	5	BU332517	603501667	750	27	96.4	931	5	BU323073	603490575
678	27	96.4	841	5	BU116154	603139121	751	27	96.4	932	7	CK180190	EST769510
679	27	96.4	841	6	CD652424	AGENCOURT	752	27	96.4	935	9	CG252438	OG0CH80TV
680	27	96.4	843	2	BF271027	GA_EB001	753	27	96.4	937	5	CG685794	OGUBT34TH
681	27	96.4	844	5	EX453505	EX453505	754	27	96.4	939	5	BU410520	603159529
682	27	96.4	844	5	CO077269	GR_Ea38P	755	27	96.4	940	5	BU109312	603110375
683	27	96.4	846	7	CK692903	ZFI01_P00	756	27	96.4	942	5	BQ682814	AGENCOURT
684	27	96.4	849	5	BU108540	603111645	757	27	96.4	942	5	CG301852	OGYBK92TH
685	27	96.4	850	7	CO004351	EST792686	758	27	96.4	944	5	AQ747405	HS_5537_A
686	27	96.4	855	7	CK800342	AGENCOURT	759	27	96.4	945	5	BU163726	AGENCOURT
687	27	96.4	855	7	CN833680	AGENCOURT	760	27	96.4	946	6	CD361912	AGENCOURT
688	27	96.4	856	7	CK775875	967014_MA	761	27	96.4	947	9	CG269007	OG4AJ59TC
689	27	96.4	860	4	BM041837	603614676	762	27	96.4	948	9	CG211948	OG3AN92TV
690	27	96.4	860	8	CC440199	PURKI72TB	763	27	96.4	949	9	CG152553	PUPUF50TB
691	27	96.4	861	6	CD102337	AGENCOURT	764	27	96.4	954	8	CC409212	PUHIQ30TB
692	27	96.4	864	2	BF341818	602016337	765	27	96.4	955	5	BU116717	603002684
693	27	96.4	864	7	CK183034	EST772349	766	27	96.4	955	8	AQ747393	HS_5537_A
694	27	96.4	864	7	CK463512	934424_MA	767	27	96.4	964	9	BZ207006	CH230-526
695	27	96.4	864	7	CN502404	AGENCOURT	768	27	96.4	964	9	CNS07021	T7_end_of
696	27	96.4	865	7	CF548740	AGENCOURT	769	27	96.4	964	9	CG055591	PUBBY57TB
697	27	96.4	865	7	CK413691	AUF_IpG11	770	27	96.4	965	1	AL534455	AL534455
698	27	96.4	867	5	BU231586	603798542	771	27	96.4	970	2	BF271758	GA_EB001
699	27	96.4	869	6	CD779573	EST650934	772	27	96.4	971	9	CNS077AN	AL432453_T3_end_of
700	27	96.4	869	7	CF820336	AGENCOURT	773	27	96.4	976	8	CC409214	PUHIQ30TD
701	27	96.4	871	6	CD327430	EST697718	774	27	96.4	977	5	BU108899	603111691
702	27	96.4	871	9	CNS045EW	Tetraodon	775	27	96.4	977	8	CK293855	EST756569
703	27	96.4	872	6	CD243432	AGENCOURT	776	27	96.4	987	7	BZ747146	PUDBY30TB
704	27	96.4	872	6	CD855556	DH0AM9ZC	777	27	96.4	988	9	CNS06YVD	T3_end_of
705	27	96.4	873	6	CA454162	AGENCOURT	778	27	96.4	993	5	CL118712	ISBI-5509
706	27	96.4	873	9	CG225540	OG2AM68TV	779	27	96.4	993	5	CL118712	ISBI-5509
707	27	96.4	874	7	CN315619	AGENCOURT	780	27	96.4	1000	5	BQ066710	AGENCOURT
708	27	96.4	875	7	CN841402	AGENCOURT	781	27	96.4	1000	5	BU120254	603144622
709	27	96.4	876	4	BG320296	Zm03_12e0	782	27	96.4	1004	9	CL427094	ZMBBBb044
710	27	96.4	877	7	CK796647	AGENCOURT	783	27	96.4	1005	6	BY706723	BY706723
711	27	96.4	878	7	CK463159	934040_MA	784	27	96.4	1007	1	AL553186	AL553186
712	27	96.4	879	9	CNS06VKK	T7_end_of	785	27	96.4	1014	9	CNS03W11	Tetraodon
713	27	96.4	880	6	CG104051	PURF23TD	786	27	96.4	1014	9	BU940736	AGENCOURT
714	27	96.4	881	6	CA493239	AGENCOURT	787	27	96.4	1019	5	BU940736	AGENCOURT
715	27	96.4	882	2	BE619446	601473276	788	27	96.4	1019	5	CL111424	ISBI-55E9
716	27	96.4	882	2	BE619446	601473276	789	27	96.4	1021	9	CC682306	OGULK61TV
717	27	96.4	884	1	AL054753	coau0001P	790	27	96.4	1023	7	CK216120	FGAS02810
718	27	96.4	885	7	CF548998	AGENCOURT	791	27	96.4	1023	6	CD255351	AGENCOURT
719	27	96.4	885	7	CG181067	602329182	792	27	96.4	1031	1	AL554066	AL554066
720	27	96.4	887	9	BX989868	Forward_8	793	27	96.4	1032	8	CC449800	ZMBBBC032
721	27	96.4	888	9	CG370741	OGYBR02TH	794	27	96.4	1035	5	BX356935	BX356935
722	27	96.4	893	6	CD302393	AGENCOURT	795	27	96.4	1040	8	BZ773311	mcv67f02_
723	27	96.4	894	7	CN501571	AGENCOURT	796	27	96.4	1047	9	CL077077	CH216-143
724	27	96.4	895	6	CD329185	AGENCOURT	797	27	96.4	1057	5	BX353739	BX353739
725	27	96.4	898	8	BZ747148	PUDBY30TD	798	27	96.4	1061	5	BX398640	BX398640
726	27	96.4	900	5	BU324673	603490174	799	27	96.4	1064	4	BM557887	AGENCOURT
727	27	96.4	902	7	CF820135	EST697517	800	27	96.4	1078	4	BM920837	AGENCOURT
728	27	96.4	902	7	CN504133	AGENCOURT	801	27	96.4	1092	5	BO712400	AGENCOURT
729	27	96.4	906	8	BZ177120	CH230-493	802	27	96.4	1094	4	BM811653	AGENCOURT
730	27	96.4	906	9	CG256135	OGXDA92TH	803	27	96.4	1102	4	BM560804	AGENCOURT
731	27	96.4	908	7	CK178004	EST767324	804	27	96.4	1105	9	CL031480	CH216-32N
732	27	96.4	908	7	CN501082	AGENCOURT	805	27	96.4	1108	8	CC274466	CH261-201
733	27	96.4	908	9	CL477515	SAIL_274	806	27	96.4	1117	9	CC597305	ZMBBBC040
734	27	96.4	909	5	BU3116876	60348241	807	27	96.4	1124	9	CNS04MRN	Tetraodon
735	27	96.4	909	5	CG211937	OG3AN92TH	808	27	96.4	1125	8	CC251818	CH261-63B
736	27	96.4	909	9	CL483192	SAIL_378	809	27	96.4	1130	9	CL082362	CH216-167
737	27	96.4	911	9	CNS04P7M	Tetraodon	810	27	96.4	1137	4	BG180651	602329424
738	27	96.4	911	9	CG252429	OG0CH80TH	811	27	96.4	1139	7	CK215408	FGAS02737
739	27	96.4	913	7	CO242078	WS00720_B	812	27	96.4	1146	6	CD498633	CDA34-F08

740	27	96.4	913	8	BZ754596	PUBEH65TD
741	27	96.4	913	9	C619302	AGUFU38TH
742	27	96.4	914	7	CN501429	AGENCOURT
743	27	96.4	914	7	CO773189	testis_ES
744	27	96.4	915	6	CD327801	AGENCOURT
745	27	96.4	916	5	BU109754	603125524
746	27	96.4	917	7	CO809863	AGENCOURT
747	27	96.4	922	2	BF346822	602021616
748	27	96.4	923	7	CN972808	20305_124
749	27	96.4	930	9	CNS021Y3	Tetraodon
750	27	96.4	931	5	BU323073	603490575
751	27	96.4	932	7	CK180190	EST769510
C 752	27	96.4	935	9	CG252438	OG0CH80TV
753	27	96.4	937	5	CG685794	OGUBT34TH
754	27	96.4	939	5	BU410520	603159529
755	27	96.4	940	5	BU109312	603110375
756	27	96.4	942	5	BQ682814	AGENCOURT
757	27	96.4	942	5	CG301852	OGYBK92TH
C 758	27	96.4	944	8	AQ747405	HS_5537_A
759	27	96.4	946	5	BU163726	AGENCOURT
760	27	96.4	946	6	CD361912	AGENCOURT
C 761	27	96.4	947	9	CG269007	OG4AJ59TC
762	27	96.4	948	9	CG211948	OG3AN92TV
763	27	96.4	949	9	CG152553	PUPUF50TB
C 764	27	96.4	954	8	CC409212	PUHIQ30TB
765	27	96.4	955	5	BU116717	603002684
C 766	27	96.4	955	8	AQ747393	HS_5537_A
767	27	96.4	958	8	BZ207006	CH230-526
C 768	27	96.4	964	9	CNS07021	T7_end_of
769	27	96.4	964	9	CG055591	PUBBY57TB
770	27	96.4	965	1	AL534455	AL534455
C 771	27	96.4	970	2	BF271758	GA_EB001
772	27	96.4	971	9	CNS077AN	AL432453 T3_end_of
773	27	96.4	976	8	CC409214	PUHIQ30TD
774	27	96.4	977	5	BU108899	603111691
775	27	96.4	979	7	CK293855	EST756569
776	27	96.4	987	8	BZ747146	PUBBY30TB
C 777	27	96.4	988	9	CNS06YVD	T3_end_of
778	27	96.4	993	6	CA980765	AGENCOURT
779	27	96.4	993	5	BU18712	603141801
C 780	27	96.4	995	9	CL11871	ISBI-5509
C 781	27	96.4	1000	5	BQ066710	AGENCOURT
C 782	27	96.4	1000	5	BU120254	603144622
C 783	27	96.4	1004	9	CL427094	ZMBBBb044
C 784	27	96.4	1005	6	BY706723	BY706723
C 785	27	96.4	1007	1	AL553186	AL553186
C 786	27	96.4	1014	9	CNS03W11	Tetraodon
C 787	27	96.4	1019	5	BU940736	AGENCOURT
C 788	27	96.4	1019	5	CL111424	ISBI-55E9
789	27	96.4	1021	9	CC682306	OGULK61TV
790	27	96.4	1023	7	CK216120	FGAS02810
791	27	96.4	1027	6	CD255351	AGENCOURT
C 792	27	96.4	1031	1	AL554066	AL554066
C 793	27	96.4	1032	8	CC449800	ZMBBBC032
794	27	96.4	1035	5	BX356935	BX356935
C 795	27	96.4	1040	8	BZ773311	mcv67f02_
796	27	96.4	1047	9	CL077077	CH216-143
797	27	96.4	1057	5	BX353739	BX353739
798	27	96.4	1061	5	CK398640	BX398640
799	27	96.4	1064	4	BN557887	AGENCOURT
800	27	96.4	1078	4	CM920837	AGENCOURT
C 801	27	96.4	1092	5	BO712400	AGENCOURT
C 802	27	96.4	1092	4	BM811553	AGENCOURT
C 803	27	96.4	1102	4	BM560804	AGENCOURT
C 804	27	96.4	1105	9	CL031480	CH216-32N
C 805	27	96.4	1108	8	CC274466	CH261-201
C 806	27	96.4	1117	9	CNS97305	ZMBBBC040
C 807	27	96.4	1124	9	CNS04WRN	Tetraodon
C 808	27	96.4	1125	8	CC251818	CH261-63B
C 809	27	96.4	1130	9	CL082362	CH216-167
C 810	27	96.4	1137	4	BG180651	CL082362
811	27	96.4	1139	7	CDK215408	FGAS02731
812	27	96.4	1146	6	CD498633	CD4314-F08

813	27	96.4	1150	4	BG2433110	BG2433110	602355211	886	26	92.9	178	1	AA339516	AA339516	EST44606
c 814	27	96.4	1158	8	C2989381	CD261-176	C 887	c 887	26	92.9	178	2	AW086987	AW086987	gal1oc08.y
c 815	27	96.4	1169	6	CD499055	CH237-B07	C 888	c 888	26	92.9	178	7	CF837447	CF837447	UCRS03.0
c 816	27	96.4	1203	8	CC254926	CH261-154	C 889	c 889	26	92.9	178	8	BH071009	RPCI-24-3	BH071009
c 817	27	96.4	1223	3	CR715567	Tetraodon	C 890	c 890	26	92.9	178	9	CG515126	OST69199	CG515126
c 818	27	96.4	1235	3	BC030742	Homo sapi	C 891	c 891	26	92.9	181	7	CK892936	SGP153004	CK892936
c 819	27	96.4	1235	9	AG332611	Mus muscu	C 892	c 892	26	92.9	181	7	CO297955	EKL1011.	CO297955
c 820	27	96.4	1327	9	CG756483	P051-3-H0	C 893	c 893	26	92.9	181	9	CR397713	Arabidops	CR397713
c 821	27	96.4	1332	9	AG441884	Mus muscu	C 894	c 894	26	92.9	182	5	BY358162	BY358162	BY358162
c 822	27	96.4	1343	4	BM921762	AGENCOURT	C 895	c 895	26	92.9	183	8	AZ779444	2M0015D22	AZ779444
c 823	27	96.4	1350	6	CD256669	AGENCOURT	C 896	c 896	26	92.9	185	8	AQ925367	RPCI-23-2	AQ925367
c 824	27	96.4	1368	8	BZ556799	Dacs1-60	C 897	c 897	26	92.9	185	3	CNS08NDX	Single re	AX019729
c 825	27	96.4	1420	9	AG591993	Mus muscu	C 898	c 898	26	92.9	187	1	AV126145	AV126145	AV126145
c 826	27	96.4	1431	9	CL105379	CL105379	C 899	c 899	26	92.9	187	8	AZ218978	Sheared D	AZ218978
c 827	27	96.4	1474	8	BZ578585	msb2 5934	C 900	c 900	26	92.9	189	5	BY244603	BY244603	BY244603
c 828	27	96.4	1477	8	AQ879012	HS_3154_A	C 901	c 901	26	92.9	193	8	AZ726200	RPCI-24-9	AZ726200
c 829	27	96.4	1521	3	CR713123	Tetraodon	C 902	c 902	26	92.9	194	5	BU075836	1m81e03.y	BU075836
c 830	27	96.4	1567	4	BG033640	602301770	C 903	c 903	26	92.9	195	8	AZ328726	1M0052F09	AZ328726
c 831	27	96.4	1625	3	AK006739	Mus muscu	C 904	c 904	26	92.9	195	8	AZ452392	1M0252B13	AZ452392
c 832	27	96.4	1690	3	AK079227	Mus muscu	C 905	c 905	26	92.9	196	2	BF952215	QV2-NN004	BF952215
c 833	27	96.4	1741	1	AF150170	AF150170	C 906	c 906	26	92.9	197	8	AZ904228	RPCI-24-1	AZ904228
c 834	27	96.4	1750	9	AG547140	Mus muscu	C 907	c 907	26	92.9	201	8	AZ052562	RPCI-23-3	AZ052562
c 835	27	96.4	1858	5	BO253383	PL (iv)79	C 908	c 908	26	92.9	202	5	BO630857	EST867281	BO630857
c 836	27	96.4	1884	3	CR591856	full-length	C 909	c 909	26	92.9	202	8	AZ889338	RPCI-24-1	AZ889338
c 837	27	96.4	2966	3	AK041235	Mus muscu	C 910	c 910	26	92.9	203	6	CA799768	sat62d06.	CA799768
c 838	27	96.4	4102	3	AK035629	Mus muscu	C 911	c 911	26	92.9	205	9	CR397712	Arabidops	CR397712
c 839	26	92.9	99	9	BX984489	Forward s	C 912	c 912	26	92.9	206	8	AZ102926	RPCI-23-2	AZ102926
c 840	26	92.9	100	1	AJ503969	AJ503969	C 913	c 913	26	92.9	207	8	AZ963112	2M0232J11	AZ963112
c 841	26	92.9	103	8	AZ666121	1M0547N21	C 914	c 914	26	92.9	208	8	AZ838565	2M0134L14	AZ838565
c 842	26	92.9	105	8	AZ270735	RPCI-23-1	C 915	c 915	26	92.9	208	8	BH229467	100615250	BH229467
c 843	26	92.9	108	8	AZ088797	RPCI-23-2	C 916	c 916	26	92.9	209	4	BG916915	602816193	BG916915
c 844	26	92.9	113	8	AZ830228	2M0109A19	C 917	c 917	26	92.9	210	9	CE678172	tigr-g88-	CE678172
c 845	26	92.9	116	8	AZ503806	1M0343B14	C 918	c 918	26	92.9	211	4	BJ041636	BJ041636	BJ041636
c 846	26	92.9	117	8	AZ774513	2M0004J05	C 919	c 919	26	92.9	211	8	AZ786353	2M0031L14	AZ786353
c 847	26	92.9	119	2	BE355781	DG1_116_A	C 920	c 920	26	92.9	212	6	CB367035	O1307 ook	CB367035
c 848	26	92.9	120	2	BF368184	CM3-CN064	C 921	c 921	26	92.9	214	2	BF048349	db81f04.y	BF048349
c 849	26	92.9	123	8	AZ062756	RPCI-23-4	C 922	c 922	26	92.9	214	5	BQ847732	Q845D08.y	BQ847732
c 850	26	92.9	124	8	AZ443186	1M0237D15	C 923	c 923	26	92.9	214	5	BQ850973	QGB14B05.	BQ850973
c 851	26	92.9	129	8	AZ299424	RPCI-23-1	C 924	c 924	26	92.9	216	2	AW844570	RC2-CN005	AW844570
c 852	26	92.9	130	2	AW936555	QV4-DT002	C 925	c 925	26	92.9	216	8	AZ092153	RPCI-23-4	AZ092153
c 853	26	92.9	131	2	BE767491	RCO-NT012	C 926	c 926	26	92.9	216	8	BZ718544	PUDAZ597B	BZ718544
c 854	26	92.9	132	8	AZ665405	1M0546P16	C 927	c 927	26	92.9	217	2	B8802389	sr29e05.y	B8802389
c 855	26	92.9	133	8	AZ773835	2M0001B10	C 928	c 928	26	92.9	217	8	AZ299282	RPCI-23-1	AZ299282
c 856	26	92.9	135	7	CK911207	e3fmjg_00	C 929	c 929	26	92.9	218	2	B8826301	QV4-EN004	B8826301
c 857	26	92.9	135	8	AZ714301	RPCI-23-3	C 930	c 930	26	92.9	219	2	BF955235	PMO-NN022	BF955235
c 858	26	92.9	137	8	AZ038342	RPCI-24-7	C 931	c 931	26	92.9	220	4	BG464863	EM1_35_D0	BG464863
c 859	26	92.9	139	8	AZ282393	RPCI-23-1	C 932	c 932	26	92.9	220	6	CB267728	100f634_H	CB267728
c 860	26	92.9	139	8	AZ649340	1M0518H17	C 933	c 933	26	92.9	221	6	CB367034	O2045 ook	CB367034
c 861	26	92.9	145	8	AZ589741	1M0398U21	C 934	c 934	26	92.9	222	1	AV323596	AV323596	AV323596
c 862	26	92.9	147	8	AZ266671	RPCI-23-1	C 935	c 935	26	92.9	222	8	AZ912231	RPCI-24-1	AZ912231
c 863	26	92.9	149	8	AZ725259	RPCI-24-9	C 936	c 936	26	92.9	225	2	AW802709	IL2-UM007	AW802709
c 864	26	92.9	150	8	AZ085617	RPCI-23-4	C 937	c 937	26	92.9	225	8	AZ223650	RPCI-23-5	AZ223650
c 865	26	92.9	150	8	BZ291320	SALK_1206	C 938	c 938	26	92.9	225	8	AZ228846	RPCI-23-8	AZ228846
c 866	26	92.9	151	2	BF172733	PCL5632_M	C 939	c 939	26	92.9	226	8	AZ303846	Forward s	AZ303846
c 867	26	92.9	153	9	CG779201	1123032B0	C 940	c 940	26	92.9	226	7	CF649942	3530_1_79	CF649942
c 868	26	92.9	157	7	W24085	zb48c08.r1	C 941	c 941	26	92.9	227	9	CG613534	OST300941	CG613534
c 869	26	92.9	157	8	AZ256887	RPCI-23-1	C 942	c 942	26	92.9	227	8	AZ656836	1M0532D12	AZ656836
c 870	26	92.9	157	8	BH104560	BH104560	C 943	c 943	26	92.9	227	9	CG513534	OST300941	CG513534
c 871	26	92.9	158	4	BG548867	EST10486	C 944	c 944	26	92.9	228	2	B8563355	EQ01f03.y	B8563355
c 872	26	92.9	166	2	BF744730	QV2-BT061	C 945	c 945	26	92.9	228	5	BQ340398	QV2-NN200	BQ340398
c 873	26	92.9	167	8	AQ981113	RPCI-23-3	C 946	c 946	26	92.9	229	1	AL810933	AL810933	AL810933
c 874	26	92.9	167	8	AZ296414	RPCI-23-1	C 947	c 947	26	92.9	230	3	CNS0AC9M	AV3335645	AV3335645
c 875	26	92.9	167	8	BH049638	RPCI-24-2	C 948	c 948	26	92.9	231	1	AV335645	Arabidops	AV335645
c 876	26	92.9	168	4	BJ574384	BJ574384	C 949	c 949	26	92.9	231	8	AQ984570	RPCI-23-3	AQ984570
c 877	26	92.9	168	8	AZ246098	RPCI-23-3	C 950	c 950	26	92.9	232	6	CB602757	M1849_M1d	CB602757
c 878	26	92.9	169	8	AZ432851	1M0218U17	C 951	c 951	26	92.9	232	8	AZ709201	RPCI-24-8	AZ709201
c 879	26	92.9	170	8	BH034935	RPCI-24-2	C 952	c 952	26	92.9	233	7	T07651	EST05541_Fe	T07651
c 880	26	92.9	172	8	AZ908235	RPCI-24-1	C 953	c 953	26	92.9	234	1	AA970041	op64d04.s	AA970041
c 881	26	92.9	172	9	CE115216	tigr-g88-	C 954	c 954	26	92.9	234	2	BF249848	pa86h10.y	BF249848
c 882	26	92.9	173	8	AZ105758	RPCI-23-4	C 955	c 955	26	92.9	234	2	CG533501	OST118821	CG533501
c 883	26	92.9	175	2	BF423559	sr37e01.y	C 956	c 956	26	92.9	235	2	B8198356	B8198356	B8198356
c 884	26	92.9	175	8	BH097797	RPCI-24-2	C 957	c 957	26	92.9	235	9	AZ707397	RPCI-23-2	AZ707397
c 885	26	92.9	176	8	BH405923	RPCI-23-1	C 958	c 958	26	92.9	235	9	CB461036	tigr-g88-	CB461036

```
c 959      26  92.9  238 7  CM670339
c 960      26  92.9  238 9  CR243557 Reverse s
c 961      26  92.9  239 8  BH405933
c 962      26  92.9  239 9  CR065268 Forward s
c 963      26  92.9  240 1  AJ461107
c 964      26  92.9  240 2  BE821729 GM700015A
c 965      26  92.9  240 6  CA599334 wawic.pk0
c 966      26  92.9  240 8  CA243330 1M0219F09
c 967      26  92.9  241 4  BI039919 MR4-NT014
c 968      26  92.9  241 4  BM503547 ih3cd07.y
c 969      26  92.9  241 8  AZ042684 RPCI-23-3
c 970      26  92.9  241 8  AZ645668 1M0511C17
c 971      26  92.9  242 2  BF961894 QV2-NN004
c 972      26  92.9  243 8  BG238903 sabs5e11.
c 973      26  92.9  243 8  AQ277514 C17B1-E1-
c 974      26  92.9  244 2  BF868877 IL5-ET011
c 975      26  92.9  244 5  BY358646 BY358646
c 976      26  92.9  244 6  CD574571 UCRPT01.0
c 977      26  92.9  244 8  AZ243508 RPCI-23-4
c 978      26  92.9  245 5  BP637160 BP637160
c 979      26  92.9  245 6  CA925130 MTU7TL.P1
c 980      26  92.9  245 7  T41759 10340.Lambd
c 981      26  92.9  246 1  AV055952
c 982      26  92.9  247 8  AZ295508 RPCI-23-1
c 983      26  92.9  247 9  CR212238 Forward s
c 984      26  92.9  248 5  BU069639 im21g04.y
c 985      26  92.9  248 8  AZ003956 RPCI-23-3
c 986      26  92.9  249 8  AZ883113 RPCI-23-2
c 987      26  92.9  250 9  BX130627 Danilo rer
c 988      26  92.9  250 9  CR228432 Reverse s
c 989      26  92.9  251 2  BB425632 BB425632
c 990      26  92.9  251 8  AZ359285 1M0102L01
c 991      26  92.9  252 1  AU063384 AU063384
c 992      26  92.9  252 8  AZ216783 1M0035D10
c 993      26  92.9  253 8  AZ272765 RPCI-23-8
c 994      26  92.9  254 4  BI395721 EST531592
c 995      26  92.9  254 8  AQ039549 CIT-HSP-2
c 996      26  92.9  254 8  AZ995886 2M0281B20
c 997      26  92.9  255 2  BE211710 S065c07.y
c 998      26  92.9  255 8  AQ934791 RPCI-23-2
c 999      26  92.9  255 8  AQ971855 RPCI-23-3
c 1000     26  92.9  255 9  CE057694 tigr-gss-

RESULT 1
BU816086/6
LOCUS      118 bp mRNA linear EST 15-OCT-2002
DEFINITION N060C10 Populus bark cDNA library Populus tremula x Populus tremuloides cDNA 5 prime, mRNA sequence.
ACCESSION  BU816086
VERSION     1
KEYWORDS    EST.
SOURCE      BU816086.1 GI:23975819
ORGANISM    Populus tremula x Populus tremuloides
            Populus tremula x Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Populus.
            1 (bases 1 to 118)
            Unneberg,P., Bhallerao,R.R., Jansson,S. and Sterky,F.
            The poplar tree transcriptome: Analysis of expressed sequence tags
            from multiple libraries
            Unpublished (2002)
            Contact: BHALLERAO RUPALI R.
            Umea Plant Science Center
            Department of Plant Physiology
            University of Umea, 901 87 Umea, Sweden
            Tel: +46 90 786 5279
            Fax: +46 90 786 6676
            Email: rupali.bhallerao@plantphys.umu.se.
            Location/Qualifiers
            1..118

FEATURES             source
ORIGIN
/organism="Populus tremula x Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:47664"
/tissue_type="bark"
/clone_lib="Populus bark cDNA library"

Alignment Scores:
Pred. No.:      847      Length:      118
Score:          27.00     Matches:      5
Percent Similarity: 83.33% Conservativeness: 0
Best Local Similarity: 83.33% Mismatches:      1
Query Match:     96.43% Indels:      0
DB:             5      Gaps:      0

US-10-030-194A-5 (1-6) x BU816086 (1-118)

Qy      1 GlyTyr***ValGluGlu 6
|||||  |||||
Db      58 GGGTATACCGTGGAGGAG 41

RESULT 2
BQ976013      131 bp mRNA linear EST 21-AUG-2002
LOCUS      QH120D04.yg.ab1 QH ABCDI sunflower RHA801 Helianthus annuus cDNA
DEFINITION clone QH120D04, mRNA sequence.
ACCESSION  BQ976013
VERSION     BQ976013.1 GI:22393536
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
            1 (bases 1 to 131)
            Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Riesberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
            Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
            Lai,Z., Church,S., Jackson,L. and Bradford,K.
            Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
            Unpublished (2002)
            Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmudson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QH_CA_Contig2725, see http://cgdb.ucdavis.edu/
            for details.
            Plate: QH120 row: D column: 04.
            Location/Qualifiers
            1..131
            /organism="Helianthus annuus"
            /mol_type="mRNA"
            /cultivar="RHA801"
            /db_xref="taxon:4232"
            /clone="QH120D04"
            /lab_host="E.coli"
            /clone_lib="QH ABCDI sunflower RHA801"
            /note="vector: pBRCNDSFIAB; The library was constructed
            from 11 different sources of RNA from a single genotype.
            Separate cDNAs were generated using primers that
            incorporated unique 5' and 3' tags to distinguish each
            source of RNA. cDNAs were then pooled, size-fractionated,
            directionally cloned into a custom medium-copy vector and
            transformations made with four size classes to minimize
            size bias. Details of each source of RNA and library
            construction can be obtained at http://cgdb.ucdavis.edu/
            TAG_LIB=QH_ABCDI sunflower environmental stress
            TAG_LIB=QH_ABCDI sunflower RHA801
```

```

ORIGIN
TAG_SEQ=CGAATGCGGG"

Alignment Scores:
Pred. No.: 958 Length: 131
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BQ978589 (1-131)

Qy 1 GlyTyr***ValGluGlu 6
Db 1 GGGTACAGTGTGGAAGAG 18

RESULT 3
BQ978589 131 bp mRNA linear EST 21-AUG-2002
DEFINITION clone QH15H23, mRNA sequence.
ACCESSION BQ978589.1 GI:22396112
VERSION BQ978589
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 131)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
JOURNAL Unpublished (2002)
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Amundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QH_CA_Contig2725, see http://cgpdb.ucdavis.edu/
for details.
Plate: QH15 row: H column: 23.
FEATURES
source
1..131
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="QH15H23"
/lab_host="E.coli"
/clone_lib="QH ABCDI sunflower RHA801"
/note="Vector: pRCNDASf1AB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=flowers environmental stress
TAG_LIB=QH ABCDI sunflower RHA801
TAG_SEQ=CGAATGCGGG"

ORIGIN
Alignment Scores:
Pred. No.: 967 Length: 132
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1

```

```

Pred. No.: 958 Length: 131
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BQ978589 (1-131)

Qy 1 GlyTyr***ValGluGlu 6
Db 1 GGGTACAGTGTGGAAGAG 18

RESULT 4
BU020428 132 bp mRNA linear EST 23-AUG-2002
LOCUS clone QHE27G13, mRNA sequence.
DEFINITION BU020428.1 GI:22455948
ACCESSION BU020428
VERSION BU020428
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 132)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
JOURNAL Unpublished (2002)
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Amundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QH_CA_Contig2725, see http://cgpdb.ucdavis.edu/
for details.
Plate: QHE27 row: G column: 13.
FEATURES
source
1..132
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHE27G13"
/lab_host="E.coli"
/clone_lib="QH EFGHJ sunflower RHA280"
/note="Vector: pRCNDASf1AB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=hulls
TAG_LIB=QH EFGHJ sunflower RHA280
TAG_SEQ=GGTAGTCGGG"

ORIGIN
Alignment Scores:
Pred. No.: 967 Length: 132
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1

```


Phasianinae; Gallus.
1 (bases 1 to 161)
Min, W., Lillehoj, H.S., Ashwell, C.M., Matukumalli, L.K., van Tassel, C. and Han, J.Y.
Chicken intestinal lymphocyte EST database as a resource for the analysis of mucosal immune function
Unpublished (2003)
Contact: Hyun S. Lillehoj
Animal Parasite Diseases Laboratory
Animal and Natural Resources Institute, USDA
Bldg. 1043, BARC-East, Beltsville, MD 20705, USA
Tel: 3015048771
Fax: 3015045103
Email: lillehoj@anri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim alt -, -trim fasta. Vector identified by cross match using options -minmatch 12 -minscore 12
Plate: 88 row: I column: 21
Seq primer: ATTAGGTGACACTATAG
High quality sequence stop: 161.
Location/Qualifiers
1..161
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="white leghorn SC"
/db_xref="taxon:9031"
/clone="IGAL 88121"
/sex="mixed"
/tissue_type="Gut"
/cell_type="Lymphocyte"
/lab_host="EMPH108"
/dev_stage="Adult"
/clone_lib="IGAL - Chicken Intestinal Lymphocyte"
/note="Organ: Intestine; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Normalized library from chicken gut infected with coccidia duodenum and middle gut."
ORIGIN
Alignment Scores:
Pred. No.: 1.22e+03 Length: 161
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0
US-10-030-194A-5 (1-6) x CD729971 (1-161)
Qy 1 GlyTyr***ValGluGlu 6
Db 57 GGGTACAGCGTGGAGGAG 74
RESULT 8
BH316542/c
LOCUS BH316542 162 bp DNA linear GSS 03-DEC-2001
DEFINITION CH230-63E6.TVB CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-63E6, genomic survey sequence.
ACCESSION BH316542
VERSION BH316542.1 GI:17247238
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 162)
Zhao, S., Shetty, J., Shatsman, S., Tsagaye, G., Geer, K., Shvartsbeyn, A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D., Riggs, P., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other GSSs: CH230-63E6.TJ
Contact: Shaying Zhao

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end page: http://www.tigr.org/tdb/bac_end3/rat/bac_end_intro.html
Plate: 63 row: E column: 6
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..162
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SeNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-63E6"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI; CHORI-230 Rat (BN/SeNHsd/MCW) BAC library produced by Pieter de Jong"
ORIGIN
Alignment Scores:
Pred. No.: 1.23e+03 Length: 162
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0
US-10-030-194A-5 (1-6) x BH316542 (1-162)
Qy 1 GlyTyr***ValGluGlu 6
Db 29 GGCTATCTGTGGAGGAA 12
RESULT 9
BU017076
LOCUS BU017076 175 bp mRNA linear EST 23-AUG-2002
DEFINITION QHE14N05.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus CDNA
clone QHE14N05, mRNA sequence.
ACCESSION BU017076
VERSION BU017076.1 GI:22452596
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.
REFERENCE
1 (bases 1 to 175)
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
Unpublished (2002)
CONTACT: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore lab
University of California at Davis (UCD)
Amundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]

belongs to contig QH_CA_Contig2725, see <http://cgpdb.ucdavis.edu/> for details.

Plate: QHE14 row: N column: 05.

FEATURES source

Location/Qualifiers
1..175
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHE14N05"
/lab_host="E.coli"
/clone_lib="QH_EFGHJ sunflower RHA280"
/note="Vector: pBRCNAsfiAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpdb.ucdavis.edu/>
TAG_TISSUE=shulls
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=GCTAGTCGGG"

ORIGIN

Alignment Scores:
Pred. No.: 1.35e+03 Length: 175
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BU017076 (1-175)

Qy 1 GlyTyr***ValGluGlu 6
||||| ||||| ||||| |||||

Db 1 GGTACAGTGTGGAGAG 18

RESULT 10 CD972375

LOCUS CD972375 175 bp mRNA linear EST 16-JUL-2003
DEFINITION QAE20h02.yg QAE Zea mays cDNA clone QAE20h02, mRNA sequence.
ACCESSION CD972375
VERSION CD972375.1 GI:32832697
KEYWORDS EST.
SOURCE Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 175)

Genoplatte.

Genoplatte, a major partnership french program in plant genomics

Unpublished (2003)

Contact: Genoplatte

Genoplatte

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french plant genome programme 'Genoplatte' (<http://www.genoplatte.com> and <http://genoplatte-info.infobiogen.fr>).

FEATURES source

Location/Qualifiers
1..175
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="F2"
/db_xref="taxon:4577"
/clone="QAE20h02"
/tissue_type="pericarp"
/clone_lib="QAE"

ORIGIN

Alignment Scores:
Pred. No.: 1.35e+03 Length: 175
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x CD972375 (1-175)

Qy 1 GlyTyr***ValGluGlu 6
||||| ||||| ||||| |||||

Db 9 GGCTACTCCGTCGAGAA 26

RESULT 11 CD972866/c

LOCUS CD972866 175 bp mRNA linear EST 16-JUL-2003
DEFINITION QAE27g02.yg QAE Zea mays cDNA clone QAE27g02, mRNA sequence.
ACCESSION CD972866
VERSION CD972866.1 GI:32833188
KEYWORDS EST.
SOURCE Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 175)

Genoplatte.

Genoplatte, a major partnership french program in plant genomics

Unpublished (2003)

Contact: Genoplatte

Genoplatte

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french plant genome programme 'Genoplatte' (<http://www.genoplatte.com> and <http://genoplatte-info.infobiogen.fr>).

FEATURES

source

Location/Qualifiers
1..175
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="F2"
/db_xref="taxon:4577"
/clone="QAE27g02"
/tissue_type="pericarp"
/clone_lib="QAE"

ORIGIN

Alignment Scores:
Pred. No.: 1.35e+03 Length: 175
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x CD972866 (1-175)

Qy 1 GlyTyr***ValGluGlu 6
||||| ||||| ||||| |||||

Db 167 GGCTACTCCGTCGAGAA 150

RESULT 12 CD976726/c

LOCUS CD976726 175 bp mRNA linear EST 16-JUL-2003
DEFINITION QAE21a03.yg QAF Zea mays cDNA clone QAE21a03, mRNA sequence.
ACCESSION CD976726
VERSION CD976726.1 GI:32837048
KEYWORDS EST.
SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

AUTHORS 1 (bases 1 to 175)

TITLE Genoplante.

JOURNAL Genoplante, a major partnership french program in plant genomics

COMMENT Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (<http://www.genoplante.com>) and <http://genoplante-info.infobiogen.fr>.

FEATURES
source Location/Qualifiers
1..175
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="F2"
/db_xref="taxon:4577"
/clone="QAF21a03"
/tissue_type="pericarp"
/clone_lib="QAF"

ORIGIN
Alignment Scores:
Pred. No.: 1.35e+03 Length: 175
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x CD976726 (1-175)

Qy 1 GlyTyr***ValGluGlu 6
|||||
167 GGCTACTCTCGTCGAAGAA 150

Db

RESULT 13
CG512665/c

LOCUS

DEFINITION CG512665 Mus musculus 129Sv/Ev Mus musculus genomic clone OST65684 genomic survey sequence.

ACCESSION CG512665

VERSION CG512665.1 GI:37299152

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 179)
Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A., Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C., Key,B.W., Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D., Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z., Sparks,M.J., Van Slightenhorst,I., Vogel,P., Walke,W., Xu,N., Zhu,Q., Person,Q. and Sands,A.T.
Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
Contact: Zambrowicz BP
OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: material@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):1608-11)
Class: Gene Trap.
Location/Qualifiers

the axillary buds. The first library, cSTA (1-20) consists of axillary buds harvested on days 1-3. This targets those genes involved in induction of the microtubers. The following libraries, cSTA (21-40) and cSTA (41-60), capture genes involved in tuber initiation and outgrowth. This library is noted as P3 in Tanksley lab notebooks."

ORIGIN

Alignment Scores:
 Pred. No.: 1.44e+03 Length: 185
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x B1177133 (1-185)

Qy 1 GlyTyr***ValGluGlu 6
 |||||
 Db 101 GGATATACAGTTGAGGAA 84

RESULT 15

B1177140/c

LOCUS

DEFINITION B1177140 185 bp mRNA linear EST 07-MAR-2003
 EST518085 cSTE Solanum tuberosum cDNA clone cSTE8H20 5' sequence,
 mRNA sequence.

ACCESSION

VERSION B1177140.1 GI:14642951

KEYWORDS EST.

SOURCE

ORGANISM

Solanum tuberosum (potato)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE

1 (bases 1 to 185)

AUTHORS

van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
 Chiemingo,A., Bougri,O., Buell,C.R., Renning,C., Tanksley,S. and
 Baker,B.

Generation of ESTs from in vitro grown microtubers

Unpublished (2001)

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr. Rockville, MD 20850, USA

Email: potato-array@tigr.org

This clone can be obtained from the University of Arizona Genomics
 Institute. Orders can be made through URL:

<http://genome.arizona.edu/orders/>

Seq primer: M13F-R.

FEATURES

Location/Qualifiers

1..185

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Bintje"

/db_xref="taxon:4113"

/clone="cSTE8H20"

/tissue_type="axillary buds of stem explants; growing

sink-tubers"

/dev_stages="7, 8 and 10 days"

/lab_host="SOLR"

/clone_lib="cSTE"

/note=Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; tissue supplied by Christian Bachem and Richard
 Visser (Department of Plant Breeding, Wageningen

University, The Netherlands). The cSTA libraries will
 attempt to capture the induction and initiation/initial

growth of the tuber in an in vitro system as described in
 Bachem et al. (Plant Journal, 1996). Small microtubers

develop from axillary buds attached to stem explants when
 placed on a high sucrose medium (10%). Visible

morphological changes occur synchronously at day five in
 the axillary buds. The first library, cSTA (1-20) consists

of axillary buds harvested on days 1-3. This targets

those genes involved in induction of the microtubers. The
 following libraries, cSTA (21-40) and cSTA (41-60),
 capture genes involved in tuber initiation and outgrowth.
 This library is noted as P3 in Tanksley lab notebooks."

ORIGIN

Alignment Scores:
 Pred. No.: 1.44e+03 Length: 185
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x B1177140 (1-185)

Qy 1 GlyTyr***ValGluGlu 6
 |||||
 Db 101 GGATATACAGTTGAGGAA 84

RESULT 16

CA904403/c

LOCUS

DEFINITION CA904403 195 bp mRNA linear EST 27-DEC-2002
 PCS13986 Scarlet Runner Bean Suspensor Region Triplex2 Phaseolus
 coccineus cDNA 5' similar to 60S ribosomal protein L24, mRNA
 sequence.

ACCESSION

VERSION CA904403.1 GI:27391395

KEYWORDS EST.

SOURCE

ORGANISM

Phaseolus coccineus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Phaseolus.

1 (bases 1 to 195)

AUTHORS

Bui,A.Q., Le,B.H., Weterings,K., Bi,Y.-P., Choi,J.-S.,
 McElroy,K.E., Choi,P.S., Harada,J.J., Fischer,R.L. and

Goldberg,R.B.

Gene Activity in Different Regions of a Post-Fertilization Plant
 Embryo by EST Analysis
 Unpublished (2002)

Contact: Goldberg, R.B.

Department of Molecular, Cell, & Developmental Biology

University of California, Los Angeles

621 Charles E. Young Drive South, Los Angeles, CA 90095-1606, USA

Tel: 310 825 3270

Fax: 310 825 8201

Email: bobg@ucla.edu

Seq primer: 5' Triplex

POLYA-No.

Location/Qualifiers

1..195

/organism="Phaseolus coccineus"

/mol_type="mRNA"

/cultivar="Hammond's Dwarf Scarlet"

/db_xref="taxon:3886"

/dev_stage="6-days post-pollination"

/clone_lib="Scarlet Runner Bean Suspensor Region Triplex2"

/note=Organ: Suspensor Region of Globular-Stage Embryos;
 Vector: Triplex2; Site 1: SfiIa; Site 2: SfiIB; Suspensor

regions were micro-dissected from globular-stage embryos
 six days after pollination from greenhouse-grown plants

[Weterings et al., Plant Cell 13, 2409-2425 (2001)].

Double-stranded cDNA was synthesized from suspensor mRNA
 using the SMART cDNA Library Construction Kit according to

the manufacturer (Clontech). The suspensor cDNA fragments
 were directionally ligated into the SfiI restriction site

of the lambda Triplex2 vector (Clontech), and the
 recombinant cDNAs were transformed into E. coli XL1-Blue

cells (Clontech). Suspensor cDNA plasmids used for
 directional sequencing were obtained by in vivo excision

from the lambda Triplex2 recombinants in E. coli BM25.8

cells (Clontech)."

```

ORIGIN
Alignment Scores:
Pred. No.: 1.53e+03 Length: 195
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x CA904403 (1-195)

QY 1 GlyTyr***ValGluGlu 6
DB 150 GGGTATACGGTTGAAGAG 133

RESULT 17
LOCUS AAS50752 201 bp mRNA linear EST 05-SEP-1997
DEFINITION nj84f09.s1 NCI_CGAP_Prl1 Homo sapiens cDNA clone IMAGE:999209, mRNA
sequence.
ACCESSION AAS50752
VERSION AAS50752.1 GI:2321004
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 201)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Rodrigo F.
Chuaqui, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 409 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 88.
FEATURES
source
1..201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:999209"
/sex="male"
/tissue_type="normal prostatic epithelial cells"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr11"
/note="Organ: prostate; Vector: pAMP10; mRNA made from
normal prostatic epithelial cells, cDNA made by oligo-dT
priming. Non-directionally cloned. Size-selected on
agarose gel, average insert size 600 bp. Library made by
D. Krizman, NIH."

ORIGIN
Alignment Scores:
Pred. No.: 1.59e+03 Length: 201
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 1 Gaps: 0

US-10-030-194A-5 (1-6) x AAS50752 (1-201)

cells (Clontech)."
```

```

QY 1 GlyTyr***ValGluGlu 6
DB 32 GGTATACTGTAGAAGAG 49

RESULT 18
LOCUS BQ912607
DEFINITION BQ912607 204 bp mRNA linear EST 19-AUG-2002
ACCESSION BQ912607
VERSION BQ912607.1 GI:22311386
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
REFERENCE 1 (bases 1 to 204)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE lettuce and sunflower ESTs from the Compositae Genome Project
JOURNAL http://compgenomics.ucdavis.edu/
COMMENT Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmunsden Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig OH_CA_Contig2725, see http://cgpdb.ucdavis.edu/
for details.
Plate: QHA21 row: H column: 01.
FEATURES
source
1..204
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="QHA21H01"
/lab_host="E.coli"
/clone_lib="OH ABCDI sunflower RHA801"
/note="Vector: pBRCNDSFIAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_SEQ=Not found"

ORIGIN
Alignment Scores:
Pred. No.: 1.61e+03 Length: 204
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BQ912607 (1-204)

QY 1 GlyTyr***ValGluGlu 6
DB 16 GGTACACTGTGAAGAG 33

RESULT 19
```

```

BB572901
LOCUS BB572901 217 bp mRNA linear EST 30-NOV-2000
DEFINITION BB572901 RIKEN full-length enriched adult male testis (DH10B) Mus
            musculus cDNA clone 4930400I20 5', mRNA sequence.
ACCESSION BB572901.1 GI:11469445
VERSION   BB572901
KEYWORDS  Mus musculus (house mouse)
SOURCE    Mus musculus
ORGANISM  Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
        1 (bases 1 to 217)
        Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T.,
        Carninci,P., Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T.,
        Hodojima,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J.,
        Kojima,Y., Konno,H., Kusakabe,M., Matsuyama,T., Miyazaki,A.,
        Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Okazaki,Y.,
        Okido,T., Owa,C., Sakai,K., Sakai,K., Sasaki,D., Sato,K.,
        Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
        Suzuki,H., Tagawa,A., Takahashi,F., Tanaka,T., Toya,T.,
        Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A.,
        Muramatsu,M. and Hayashizaki,Y.
        RIKEN Mouse ESTs (Aizawa,K. et al. 2000)
        Unpublished (2000)
        Contact: Yoshihide Hayashizaki
        Laboratory for Genome Exploration Research Group, RIKEN Genomic
        Sciences Center(GSC), Yokohama Institute
        The Institute of Physical and Chemical Research (RIKEN)
        1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
        Tel: 81-45-503-9222
        Fax: 81-45-503-9216
        Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
        Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
        Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
        Thermotabilization and thermoactivation of thermostable enzymes by
        trehalose and its application for the synthesis of full length
        cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
        Itoh,M., Kiteunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
        Tonaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
        Okazaki,Y. and Hayashizaki,Y.
        Automated filtration-based high-throughput plasmid preparation
        system. Genome Res. 9 (5), 463-470 (1999)
        Carninci,P. and Hayashizaki,Y.
        High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
        19-44 (1999)
        Please visit our web site (http://genome.rtc.riken.go.jp) for
        further details.
FEATURES             source
        source
        1..217
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="CS7BL/6J"
        /db_xref="taxon:10090"
        /clone="4930400I20"
        /sex="male"
        /tissue_type="testis"
        /dev_stage="adult"
        /lab_host="DH10B"
        /clone_lib="RIKEN full-length enriched, adult male testis
        (DH10B)"
        /note="Site 1: SalI; Site 2: BamHI; cDNA library was
        prepared and sequenced in Mouse Genome Encyclopedia
        Project of Genome Exploration Research Group in Riken
        Genomic Sciences Center and Genome Science Laboratory in
        RIKEN. Division of Experimental Animal Research in Riken
        contributed to prepare mouse tissues. 1st strand cDNA was
        primed with a primer [5'
        GAGAGAGAGAGATCCCAAGACTCTTTTTTTTTTTTTT 3']. cDNA was
        prepared by using trehalose thermo-activated reverse
        transcriptase and subsequently enriched for full-length by
        cap-trapper. Second strand cDNA was prepared with the
        primer adapter of sequence [5'
        GAGAGAGAGATTCGAGTTAATTAATAATTCCTCCCCCCCCCCC 3']. cDNA

```

```

ORIGIN
Alignment Scores:
Pred. No.: 1.74e+03 Length: 217
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 2 Gaps: 0

US-10-030-194a-5 (1-6) x BB572901 (1-217)

QY 1 GlyTyr***ValGluGlu 6
Db 82 GGCTACTCAGTTGAAGAA 99
      ||||| ||||| |||||
      ||||| ||||| |||||

RESULT 20
LOCUS CB129746/C 217 bp mRNA linear EST 29-JAN-2003
DEFINITION K-EST0179471 L11SNU354 Homo sapiens cDNA clone L11SNU354-6-H02 5',
            mRNA sequence.
ACCESSION CB129746
VERSION   CB129746.1 GI:28093546
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
        1 (bases 1 to 217)
        Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
        Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
        Kim,Y.S.
        21C Frontier Korean EST Project 2001
        Unpublished (2002)
        Contact: Kim YS
        Genome Research Center
        Korea Research Institute of Bioscience & Biotechnology
        52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea
        Tel: +82-42-860-4470
        Fax: +82-42-860-4409
        Email: yongsung@mail.kribb.re.kr
        Plate: 6 row: H column: 02
        High quality sequence stop: 217.
FEATURES             Location/Qualifiers
        source
        1..217
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="L11SNU354-6-H02"
        /sex="M"
        /tissue_type="Liver"
        /cell_type="Polygonal"
        /cell_line="SNU-354"
        /lab_host="Top10P"
        /clone_lib="L11SNU354"
        /note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
        Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
        bacterial alkaline phosphatase (BAP) and then decapped
        with tobacco acid pyrophosphatase (TAP). The decapped
        intact mRNA was ligated with DNA-RNA linker including
        EcoRI site by treatment of T4 RNA ligase and the first
        strand cDNA was synthesized from oligo dt-selected mRNA by
        priming with dt-tailed vector. The dt-tailed vector was
        adjusted to have about 60nt. The cDNA vector was
        circularized with E. coli DNA ligase after digestion of
        EcoRI which site is also included in vector. An RNA strand
        converted to a DNA strand by Okayama-Berg method. The
        obtained cDNA vectors were used for transformation of
        competent cells E. coli Top10P by electroporation method.
        The cDNA libraries constructed by this method are

```

was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."

full-length enriched cDNA library."

```

ORIGIN
Alignment Scores:
Pred. No.:      1.74e+03      Length:      217
Score:          27.00      Matches:      5
Percent Similarity: 83.33%      Conservative: 0
Best Local Similarity: 83.33%      Mismatches: 1
Query Match:    96.43%      Indels:    0
DB:             6      Gaps:          0

US-10-030-194A-5 (1-6) x CB129746 (1-217)

QY      1 GlyTyr***ValGluGlu 6
|||||  |||||  |||||  |||||  |||||
Db      125 GGTATGTCAGTGGGAAG 108

RESULT 21
BU023155
LOCUS   BU023155
DEFINITION   BU023155 218 bp mRNA linear EST 23-AUG-2002
              clone QHE9L24, mRNA sequence.
ACCESSION   BU023155
VERSION     BU023155.1 GI:22458675
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
ORGANISM    Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
REFERENCE   1 (bases 1 to 218)
AUTHORS    Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
            Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
            Lai,Z., Church,S., Jackson,L. and Bradford,K.
            Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
            Unpublished (2002)
JOURNAL
COMMENT    Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmunden Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QH_CA_Contig2725, see http://cgpdb.ucdavis.edu/
            for details.
Plate: QHE9 row: L column: 24.
FEATURES             Location/Qualifiers
     source           1..218
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA280"
                     /db_xref="taxon:4232"
                     /clone="QHE9L24"
                     /lab_host="E.coli"
                     /clone_lib="QH_EFGHJ sunflower RHA280"
                     /note="Vector: pBRCDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_TISSUE=shoots environmental stress
                     TAG_LIB=QH_EFGHJ sunflower RHA280
                     TAG_SEQ=TCGCAACGGG"

ORIGIN
Alignment Scores:

```

```

Pred. No.:      1.75e+03      Length:      218
Score:          27.00      Matches:      5
Percent Similarity: 83.33%      Conservative: 0
Best Local Similarity: 83.33%      Mismatches: 1
Query Match:    96.43%      Indels:    0
DB:             5      Gaps:          0

US-10-030-194A-5 (1-6) x BU023155 (1-218)

QY      1 GlyTyr***ValGluGlu 6
|||||  |||||  |||||  |||||  |||||
Db      96 GGTACAGTGTGGAAG 113

RESULT 22
BF509543
LOCUS   BF509543
DEFINITION   BF509543 221 bp mRNA linear EST 06-DEC-2000
              UI-H-B14-a0z-h-04-0-UI-s1 NCI CGAP_Sub8 Homo sapiens cDNA clone
              IMAGE:3086815 3', mRNA sequence.
ACCESSION   BF509543
VERSION     BF509543.1 GI:11592841
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 221)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
JOURNAL
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. cDNA library preparation: M.B. Soares Lab Clone distribution:
            NCI-CGAP clone distribution information can be found through the
            I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seg primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source           1..221
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3086815"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI-CGAP Sub8"
                     /note="Vector: p7T73D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: Not 1; Site 2: Eco RI; NCI CGAP Sub8
                     is a subtracted library derived from NCI CGAP Sub5. The
                     NCI CGAP Sub8 library had 2.5 million recombinants. A
                     single-stranded DNA preparation of NCI-CGAP Sub5 was used
                     as a tracer in a subtractive hybridization with a driver
                     comprising: a pool of clones from NCI CGAP Sub5 (IMAGE
                     clone ids 2732833-2737415, 3068040-3069191; 25% of the
                     driver population), a pool of clones from NCI CGAP Sub4
                     (IMAGE clone ids 2723592-2729326; 25% of the driver
                     population), NCI-CGAP Sub6 (pool AIF-AJU, IMAGE ids
                     2728969-2733190; 25% of the driver population), and
                     NCI CGAP Sub7 (IMAGE ids 3069192-3072238,
                     3081864-3084550; 25% of the driver population).
                     Subtraction was performed as previously described
                     [Bonaldi, Lennon & Soares (1996): Normalization and
                     Subtraction: Two Approaches to Facilitate Gene Discovery.
                     Genome Research 6, 791-806.
                     TAG_TISSUE=germ cell
                     TAG_LIB=NCI CGAP_GC4
                     TAG_SEQ=AAATC"

ORIGIN
Alignment Scores:

```

```

Pred. No.: 1.77e+03 Length: 221
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-5 (1-6) x BF509543 (1-221)

Qy 1 GlyTyr***ValGluGlu 6
Db 188 GGTATTAGTGTGGAAGAG 205

RESULT 23
BU020431 228 bp mRNA linear EST 23-AUG-2002
LOCUS clone QHE27G17, mRNA sequence.
DEFINITION BU020431
ACCESSION BU020431.1 GI:22455951
VERSION EST
KEYWORDS Helianthus annuus (common sunflower)
SOURCE Helianthus annuus
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 228)
AUTHORS Kozik,A., Micheltore,R.W., Knapp,S., Matvienko,M., Riesberg,L.,
Lin,H., van Damm,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE Lettuce and sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
JOURNAL Unpublished (2002)
COMMENT Contact: Alexander Kozik [R.W.Micheltore]
Department of Vegetable Crops, R.W.Micheltore Lab
University of California at Davis (UCD)
Amundson Hall, UCD Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [micheltore@vegmail.ucdavis.edu]
belongs to contig QH_CA_Contig2725, see http://cgdb.ucdavis.edu/
for details.
Plate: QHE27 row: G column: 17.

FEATURES
source
1..228
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHE27G17"
/lab_host="E.coli"
/clone_lib="QH_EFGHJ sunflower RHA280"
/notes="Vector: pBRCDNASFIAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgdb.ucdavis.edu/
TAG_TISSUE=shoots environmental stress
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCAACGGG"

ORIGIN
Alignment Scores:
Pred. No.: 1.84e+03 Length: 228
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-5 (1-6) x BF509543 (1-221)

Qy 1 GlyTyr***ValGluGlu 6
Db 188 GGTATTAGTGTGGAAGAG 205

RESULT 24
BU020431 231 bp DNA linear GSS 10-OCT-2002
LOCUS clone QHE27G17, genomic survey sequence.
DEFINITION BU020431
ACCESSION BU020431.1 GI:23726375
VERSION GSS
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 231)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Teagaye,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,S., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
JOURNAL Other_GSSs: CH230-214L3.TV
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orering\_information.htm). BAC end
page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html
Plate: 214 row: L column: 3
Seq primer: SP6
Class: BAC ends.

FEATURES
source
1..231
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-214L3"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/notes="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
Alignment Scores:
Pred. No.: 1.87e+03 Length: 231
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x BZ090362 (1-231)

Qy 1 GlyTyr***ValGluGlu 6
Db 163 GGTATTAGCGTTGAGGAG 146

```

```

RESULT 25
BU018501
LOCUS
DEFINITION
OHE18N13.yg.ab1 OH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
clone OHE18N13, mRNA sequence.
ACCESSION
BU018501
VERSION
BU018501.1 GI:22454021
KEYWORDS
SOURCE
ORGANISM
Helianthus annuus (common sunflower)
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE
1 (bases 1 to 232)
Kozik,A., Micheltore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
JOURNAL
COMMENT
Contact: Alexander Kozik [R.W.Micheltore]
Department of Vegetable Crops, R.W.Micheltore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [micheltore@vegmail.ucdavis.edu]
belongs to contig OH_CA_Contig2725, see http://cgpdb.ucdavis.edu/
for details.
Plate: OHE18 row: N column: 13.
FEATURES
source
Location/Qualifiers
1..232
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="OHE18N13"
/lab_host="E.coli"
/clone_lib="OH_EFGHJ sunflower RHA280"
/note="Vector: pBRCDNASFIAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=shoots environmental stress
TAG_LIB=OH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCAACGGG"
ORIGIN
Alignment Scores:
Pred. No.: 1..88e+03 Length: 232
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0
US-10-030-194A-5 (1-6) x BU018501 (1-232)
Qy 1 GlyTyr***ValGluGlu 6
| | | | |
Db 102 GGGTACAGTGTGGAAGAG 119
| | | | |
RESULT 26
BQ968398
LOCUS
DEFINITION
OHE18N13.yg.ab1 OH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
clone OHE18N13, mRNA sequence.
ACCESSION
BQ968398
VERSION
BQ968398.1 GI:22385919
KEYWORDS
SOURCE
ORGANISM
Helianthus annuus (common sunflower)
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE
1 (bases 1 to 239)
Kozik,A., Micheltore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
JOURNAL
COMMENT
Contact: Alexander Kozik [R.W.Micheltore]
Department of Vegetable Crops, R.W.Micheltore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [micheltore@vegmail.ucdavis.edu]
belongs to contig OH_CA_Contig2725, see http://cgpdb.ucdavis.edu/
for details.
Plate: QHE18 row: M column: 10.
FEATURES
source
Location/Qualifiers
1..239
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="QHE18N10"
/lab_host="E.coli"
/clone_lib="OH_ABCDI sunflower_RHA801"
/note="Vector: pBRCDNASFIAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=chemical induction
TAG_LIB=OH_ABCDI sunflower RHA801
TAG_SEQ=TCGTAGCCGGG"
ORIGIN
Alignment Scores:
Pred. No.: 1..95e+03 Length: 239
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0
US-10-030-194A-5 (1-6) x BQ968398 (1-239)
Qy 1 GlyTyr***ValGluGlu 6
| | | | |
Db 1 GGGTACAGTGTGGAAGAG 18
| | | | |
RESULT 27
AA301321
LOCUS
DEFINITION
EST14252 Testis tumor Homo sapiens cDNA 5' end similar to similar
to polypyrimidine tract-binding protein PTB, mRNA sequence.
ACCESSION
AA301321
VERSION
AA301321.1 GI:1953654

```

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@igr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
FEATURES
source
1. .244
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):191684"
/db_xref="taxon:9606"
/sex="male"
/dev_stage="adult"
/clone_lib="Testis tumor"
/note="Organ: testis; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
ORIGIN
Alignment Scores:
Pred. No.: 1.99e+03 Length: 244
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 1 Gaps: 0
US-10-030-194A-5 (1-6) x AA301321 (1-244)
QY
1 GlyTyr***ValGluGlu 6
|||||
Db
14 GGATATTCTGTAGAAGAA 31
RESULT 28
AW477834/c AW477834 247 bp mRNA linear EST 09-JUL-2000
LOCUS
DEFINITION
17053 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION
AW477834
VERSION
AW477834.1 GI:7047940
KEYWORDS
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 244)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,J.C., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,F., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dinke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
7566098
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@igr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
FEATURES
source
1. .247
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/note="vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
ORIGIN
Alignment Scores:
Pred. No.: 2.02e+03 Length: 247
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 2 Gaps: 0
US-10-030-194A-5 (1-6) x AW477834 (1-247)
QY
1 GlyTyr***ValGluGlu 6
|||||
Db
78 GGGTACTCTGTGGAGGAG 61
RESULT 29
BZ260573 261 bp DNA linear GSS 15-OCT-2002
BZ260573
LOCUS
DEFINITION
CH230-285M19.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-285M19, genomic survey sequence.
ACCESSION
BZ260573
VERSION
BZ260573.1 GI:23969998
KEYWORDS
GSS.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 261)
Zhao,S., Shetty,J., Shatsman,S., Teegave,G., Geer,K.,
Shivatsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
JOURNAL

```

```

SOURCE
ORGANISM
Sus scrofa (pig)
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 247)
Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Pertes,G., Sultana,R.,
Quackenbush,J. and Keele,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
22213789
MEDLINE
PUBMED
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 7 row: N column: 20
Seq primer: ATTTAGGTGACACTATAG.
FEATURES
source
1. .247
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/note="vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
ORIGIN
Alignment Scores:
Pred. No.: 2.02e+03 Length: 247
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 2 Gaps: 0
US-10-030-194A-5 (1-6) x AW477834 (1-247)
QY
1 GlyTyr***ValGluGlu 6
|||||
Db
78 GGGTACTCTGTGGAGGAG 61
RESULT 29
BZ260573 261 bp DNA linear GSS 15-OCT-2002
BZ260573
LOCUS
DEFINITION
CH230-285M19.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-285M19, genomic survey sequence.
ACCESSION
BZ260573
VERSION
BZ260573.1 GI:23969998
KEYWORDS
GSS.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 261)
Zhao,S., Shetty,J., Shatsman,S., Teegave,G., Geer,K.,
Shivatsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
JOURNAL

```

COMMENT

Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.choi.org).
Clones may be purchased from BACPAC Resources
(<http://www.chori.org/bacpac/orering/information.html>). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 285 row: M column: 19
Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..261
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-285M19"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Alignment Scores:

Pred. No.: 2.16e+03 Length: 261
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-5 (1-6) x BZ260573 (1-261)

QY 1 GlyTyr***ValGluGlu 6

|||||TCTTTAGAGGAA 61
DB 44 GCGTATTCTGTAGAGGAA 61

RESULT 30

CG505068/c

LOCUS

CG505068 261 bp DNA linear GSS 01-OCT-2003
OST53833 Mus musculus 129Sv/Ev Mus musculus genomic clone OST53833,
genomic survey sequence.

ACCESSION

CG505068

VERSION

GSS.

KEYWORDS

SOURCE

Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 261)
Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
Piggott,J., BeltrandeRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
Key,B.W. Jr., Kipp,P., Kohlhaufl,B., Ma,Z.-Q., Markesich,D.,
Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
Zhu,Q., Person,C. and Sands,A.T.
Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
Contact: Zambrowicz Bp
OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com

FEATURES

source

Location/Qualifiers

1..261
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"
/clone="OST53833"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

ORIGIN

Alignment Scores:

Pred. No.: 2.16e+03 Length: 261
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-5 (1-6) x CG505068 (1-261)

QY 1 GlyTyr***ValGluGlu 6

|||||TCTTTAGAGGAG 138
DB 155 GCGTATTCTGTGAAGAG 138

RESULT 31

CG521180

LOCUS

CG521180 262 bp mRNA linear EST 07-JUL-2004
LIONP462G07101 3', mRNA sequence.

ACCESSION

CG521180

VERSION

EST.

KEYWORDS

SOURCE

Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 262)
Henrich,J., Hermanns,J., Kranz,H., Loebbert,R., Schlueter,T.,
Schuette,D., Weindel,M., Heil,O., Ebert,L., Neubert,P., Peters,M.,
Radelof,U., Schneider,D. and Korn,B.
Mouse ArrayTAG cDNA
Unpublished (2004)
Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; LIONP462G07101.
RZPDLIB;
Mouse ArrayTAG cDNA
<http://www.rzpd.de/cgi-bin/products/showlib.pl.cgi?response?libNo=462> Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de

This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
RP: CAGGAACAGCATGAC.

FEATURES

source

Location/Qualifiers

1..262
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="LIONP462G07101"
/lab_host="DH10B"
/clone_lib="Mouse pBluescript Lion"

ORIGIN

```

Alignment Scores:
Pred. No.: 2.17e+03 Length: 262
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 83.33% Indels: 1
DB: 7 Gaps: 0

US-10-030-194A-5 (1-6) x CR521180 (1-262)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||
Db 150 GCCTATGCTGTGGAAGAG 167

RESULT 32
AA334794 264 bp mRNA linear EST 21-APR-1997
LOCUS EST39059 Embryo, 9 week Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA334794
VERSION AA334794.1 GI:1987037
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 264)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glocke,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1..264
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):136480"
/db_xref="taxon:9606"
/dev_stage="embryo, 9 wks"
/clone_lib="Embryo, 9 week"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 2.19e+03 Length: 264
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 83.33% Indels: 1
DB: 7 Gaps: 0

US-10-030-194A-5 (1-6) x CR521180 (1-262)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||
Db 150 GCCTATGCTGTGGAAGAG 167

RESULT 32
AA334794 264 bp mRNA linear EST 21-APR-1997
LOCUS EST39059 Embryo, 9 week Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA334794
VERSION AA334794.1 GI:1987037
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 264)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glocke,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1..264
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):136480"
/db_xref="taxon:9606"
/dev_stage="embryo, 9 wks"
/clone_lib="Embryo, 9 week"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 2.19e+03 Length: 264
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 83.33% Indels: 1
DB: 7 Gaps: 0

US-10-030-194A-5 (1-6) x AA334794 (1-264)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||
Db 108 GCATATGCAGTAGAGAA 125

RESULT 33
AG228772/c
LOCUS AG228772 270 bp DNA linear GSS 19-JUL-2003
DEFINITION Lotus corniculatus var. japonicus DNA, clone:LjB16006_r, genomic
survey sequence.
ACCESSION AG228772
VERSION AG228772.1 GI:26539396
KEYWORDS GSS.
SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
Lotus.
REFERENCE 1
AUTHORS Sato,S., Nakamura,Y. and Tabata,S.
TITLE Lotus japonicus BAC End sequences
JOURNAL Published Only in Database (2002)
REFERENCE 2 (bases 1 to 270)
AUTHORS Sato,S.
TITLE Direct Submision
JOURNAL Submitted (20-NOV-2002) Shusei Sato, Kazusa DNA Research Institute,
The First Laboratory for Plant Gene Research; 2-6-7
Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:ssato@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/,
Tel:81-438-52-3935(ex.2336), Fax:81-438-52-3934)
FEATURES
source
1..270
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/strain="Miyakojima MG-20"
/variety="japonicus"
/db_xref="taxon:34305"
/clone="LjB16006_r"
/clone_lib="genomic BAC library"
/note="VECTOR:pBelOBAC11-synonym: Lotus japonicus"

ORIGIN
Alignment Scores:
Pred. No.: 2.25e+03 Length: 270
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-5 (1-6) x AG228772 (1-270)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||
Db 219 GCATATGCAGTAGAGAG 202

RESULT 34
AV278603/c
LOCUS AV278603 271 bp mRNA linear EST 05-NOV-1999
DEFINITION AV278603 RIKEN full-length enriched, adult male testis (DH10B) Mus
musculus cDNA clone 4933402N03 3', mRNA sequence.
ACCESSION AV278603
VERSION AV278603.1 GI:6266640
KEYWORDS Mus musculus (house mouse)
SOURCE

```


TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
<http://compgenomics.ucdavis.edu/>

JOURNAL Unpublished (2002)

COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmudson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozika@ucdavis.org [michelmore@vegmail.ucdavis.edu]
belongs to contig OH_CA_Contig275, see <http://cgdb.ucdavis.edu/>
for details.
Plate: QHAl6 row: G column: 21.

FEATURES
source Location/Qualifiers
1..281
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RH801"
/db_xref="taxon:4232"
/cloned="QHAl6G21"
/lab_host="E.coli"
/cloned_lib="OH ABCDI sunflower RH801"
/note="Vector: pPERCNDASfiAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgdb.ucdavis.edu/>
TAG_SEQ=Not found"

ORIGIN

Alignment Scores:			
Pred. No.:	2,35e+03	Length:	281
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	5	Gaps:	0

US-10-030-194A-5 (1-6) x BQ911194 (1-281)

QY 1 GlyTyr***ValgluGlu 6

Db 29 GGTTACAGTGCGAAGAG 46

RESULT 38
CG502243/c

LOCUS CG502243 Mus musculus 129SV/Ev Mus musculus genomic clone OS746996 linear GSS 01-OCT-2000

DEFINITION genomic survey sequence.

ACCESSION CG502243

VERSION CG502243.1 GI:37276785

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;

REFERENCE 1 (bases 1 to 284)

AUTHORS Zambronicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J.,
Piggott, J.C., BeltrandelRio, H., Buxton, E.C., Edwards, J., Finch, R.A.,
Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C.,
Key, B.W. Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D.,
Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,
Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N.,
Zhu, Q., Person, C. and Sands, A.T.

TITLE Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

COMMENT Contact: Zambrowicz BP
OmniBank

Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676): 608-11)
Class: Gene Trap.

FEATURES
source Location/Qualifiers
1. .284
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"
/clone="OST46996"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

ORIGIN

Alignment Scores:	2.38e+03	Length:	284
Pred. No.:	27.00	Matches:	5
Score:	83.33%	Conservative:	0
Percent Similarity:	83.33%	Mismatches:	1
Best Local Similarity:	96.43%	Indels:	0
Query Match:	9	Gaps:	0

DB:

US-10-030-194A-5 (1-6) x CG502243 (1-284)

Qy 1 GlyTyr***ValGluGlu 6

Db 238 GCGTATGCTGTGGAAAG 221

RESULT 39

BU019672

LOCUS

DEFINITION BU019672 287 bp mRNA linear EST 23-AUG-2002
QHE22M17.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus CDNA

ACCESSION

VERSION BU019672

KEYWORDS EST.

SOURCE Helianthus annuus (common sunflower)

ORGANISM Helianthus annuus

REFERENCE 1 (bases 1 to 287)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Riesberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://comgenomics.ucdavis.edu/
Unpublished (2002)

AUTHORS Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Axmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659

TITLE Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QH_CA_Contig2725, see http://cgpdb.ucdavis.edu/
for details.

JOURNAL Plate: QHE22 row: M column: 17.

COMMENT Location/Qualifiers
1. .287
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHE22M17"
/lab_host="E.coli"
/clone_lib="QH_EFGHJ sunflower RHA280"
/note="Vector: pBRCNAB51AB; The library was constructed

from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at <http://cgpdb.ucdavis.edu/>
TAG_TISSUE=hulls
TAG_LIB=QH_EFGHJ_sunflower_RHA280
TAG_SEQ=GCTAGTCGGG"

ORIGIN

Alignment Scores:
Pred. No.: 2.41e+03 Length: 287
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BU019672 (1-287)

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 1 GGGTACACTGTGGAAGAG 18

RESULT 40

CF034190 287 bp mRNA linear EST 17-JUL-2003
LOCUS CF034190 QCF Zea mays cDNA clone QCF2c02, mRNA sequence.

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 287)
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.inbio.gen.fr>).

FEATURES

source
1..287
/organism="Zea mays"
/mol_type="mRNA"
/culivar="F2"
/db_xref="taxon:4577"
/clone="QCF2c02"
/tisue type="seedling minus kernel"
/clone_lib="QCF"

ORIGIN

Alignment Scores:
Pred. No.: 2.41e+03 Length: 287
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x CF034190 (1-287)

Qy 1 GlyTyr***ValGluGlu 6

Db 20 GGCTACTCTCGTGAAGAA 37
|||||
|||||

RESULT 41

CG502254/c 287 bp DNA linear GSS 01-OCT-2003
LOCUS CG502254/c Mus musculus 129Sv/Ev Mus musculus genomic clone OST47010,
DEFINITION genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 287)
Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J.,
Piggott, J., BeltrandeRio, H., Buxton, E.C., Edwards, J., Finch, R.A.,
Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C.,
Key, B.W. Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D.,
Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,
Spark, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N.,
Zhu, Q., Person, C. and Sands, A.T.
Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
Contact: Zambrowicz BP
OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene trap.
Location/Qualifiers
1..287
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"
/clone="OST47010"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

TITLE

JOURNAL

COMMENT

FEATURES

source
1..287
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"
/clone="OST47010"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

ORIGIN

Alignment Scores:
Pred. NO.: 2.41e+03 Length: 287
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-5 (1-6) x CG502254 (1-287)

Qy 1 GlyTyr***ValGluGlu 6

Db 241 GGCTATCTCTGTGGAAGAG 224
|||||
|||||

RESULT 42

AV144507/c 289 bp mRNA linear EST 03-JUL-1999
LOCUS AV144507/c Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA
DEFINITION clone 2810437E14, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 289)
Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,

Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tomimaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muranatsu, M., Okazaki, Y. and Hayashizaki, Y.

RIKEN Mouse ESTs
Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp
Thermotabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source
1..289
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2810437E14"
/sex="mixed"
/dev_stage="10-11 day embryo"
/clone_lib="Mus musculus C57BL/6J 10-11 day embryo"

ORIGIN

Alignment Scores:
Pred. No.: 2,438+03 Length: 289
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 1 Gaps: 0

US-10-030-194A-5 (1-6) x AV144507 (1-289)

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 131 GGTACACAGTTGAGAA 114

RESULT 43

LOCUS BI046607 293 bp mRNA linear EST 14-JUN-2001
DEFINITION MR3-FN0206-070201-013-d08 FN0206 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI046607
VERSION BI046607.1 GI:14453229
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 293)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR3&t2=MR3-FN0206-070201-013-d08&t3=2001-02-07&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 43
High quality sequence stop: 292.

FEATURES

source
1..293
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="FN0206"

/note="Organ: prostate normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:
Pred. No.: 2,478+03 Length: 293
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x BI046607 (1-293)

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 23 GGGTATCGGTTGAGAA 40

RESULT 44

LOCUS CK248536 295 bp mRNA linear EST 30-JUL-2004
DEFINITION EST732173 potato callus cDNA library, normalized and full-length Solanum tuberosum cDNA clone POCB168 3' end, mRNA sequence.

ACCESSION CK248536
VERSION CK248536.1 GI:39798767
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)

ORGANISM

Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 295)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

REFERENCE

AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.
Seq primer: GTA ATA CGA CTC ATA GGG C.

FEATURES

source
1..295
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"

/db_xref="taxon:4113"
 /clone="POCB168"
 /tissue_type="callus"
 /lab_host="DH10B-TonA"
 /clone_lib="potato callus cDNA library, normalized and full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Solanum tuberosum var. Kennebec callus tissue grown on solid media."

ORIGIN

Alignment Scores:
 Pred. No.: 2.49e+03 Length: 295
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 7 Gaps: 0

US-10-030-194A-5 (1-6) x CK248536 (1-295)

Qy 1 GlyTyr***ValGluGlu 6
 ||||| |||||
 Db 213 GGATATACAGTTGAGAA 230

RESULT 45
 AW086727/c
 LOCUS AW086727 298 bp mRNA linear EST 22-OCT-1999
 DEFINITION ga04e02.x1 Moss EST library CPU Ceratodon purpureus cDNA clone
 PEP SOURCE ID:CPU011503 3' similar to TR:082061 082061 PROTEIN INVOLVED IN STARCH METABOLISM PRECURSOR. ;, mRNA sequence.

ACCESSION AW086727
 VERSION AW086727.1 GI:6042430
 KEYWORDS EST.
 SOURCE Ceratodon purpureus

ORGANISM Ceratodon purpureus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Dicranidae; Dicranales; Ditrichaceae; Ceratodon.

REFERENCE 1 (bases 1 to 298)
 AUTHORS Quatrano, R., Bashardes, S., Cove, D., Cuming, A., Knight, C., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Steptoe, M., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.
 TITLE Leeds/Wash U Moss EST Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Ralph Quatrano
 Leeds/Wash U Moss EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Libraries were constructed by Dr. Stavros Bashardes as part of the Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and Washington Univ. in St. Louis (USA) DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
 Seq primer: -400P from Gibco.
 Location/Qualifiers
 1..298
 /organism="Ceratodon purpureus"
 /db_type="mRNA"
 /db_xref="taxon:3225"
 /clone="PEP SOURCE ID:CPU011503"
 /tissue_type="protonemata: 7 day old tissue"
 /lab_host="DH10B"
 /clone_lib="Moss EST library CPU"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Construction of the cDNA library was carried out using Stratagene's UniZAP - cDNA synthesis kit. cDNA was constructed using an oligo dT primer/linker that contains a XhoI site within it. Following ds cDNA synthesis, EcoRI

FEATURES
 source

1..298
 /organism="Ceratodon purpureus"
 /db_type="mRNA"
 /db_xref="taxon:3225"
 /clone="PEP SOURCE ID:CPU011503"
 /tissue_type="protonemata: 7 day old tissue"
 /lab_host="DH10B"
 /clone_lib="Moss EST library CPU"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Construction of the cDNA library was carried out using Stratagene's UniZAP - cDNA synthesis kit. cDNA was constructed using an oligo dT primer/linker that contains a XhoI site within it. Following ds cDNA synthesis, EcoRI

adapters were ligated to the blunt ends and sample was digested with XhoI. The result is cDNA with an EcoRI sticky end on one side and a XhoI sticky end on the other. This cDNA was ligated directionally in UniZAP arms. The vector is designed containing the pBluescript sequence as well as lambda DNA and cDNA is cloned within this pBluescript sequence. The vector was then packaged using Gold gigapackaging extracts. Library was grown in XL1Blue MR' cells and amplified. The library was excised by mass excision using Stratagene's Mass excision kit, that uses exassist as a helper phage that releases the pBluescript sequence and circularises it as single stranded plasmids that are then packaged (by helper phage) and secreted out of the host cell as phagemids. SOLR cells were transformed with phagemids and the library was plated out on LB-amp plates to select for transformants. Approximately 1,000,000 colonies were grown and recovered. The double stranded plasmid library was recovered by using Qiaagen Midi prep kit. 2 micro grams of each library were used to transform DH10B cells by electroporation."

ORIGIN

Alignment Scores:
 Pred. No.: 2.52e+03 Length: 298
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 2 Gaps: 0

US-10-030-194A-5 (1-6) x AW086727 (1-298)

Qy 1 GlyTyr***ValGluGlu 6
 ||||| |||||
 Db 238 GGATACGCTGTCGAGAG 221

RESULT 46

CA516999
 LOCUS CA516999 298 bp mRNA linear EST 15-NOV-2002
 DEFINITION KS09068E11 KS09 Capsicum annum cDNA, mRNA sequence.
 ACCESSION CA516999
 VERSION CA516999.1 GI:25017556
 KEYWORDS EST.
 SOURCE Capsicum annum

ORGANISM Capsicum annum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Capsicum.
 REFERENCE 1 (bases 1 to 298)
 AUTHORS Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S., Hur, C.-G. and Choi, D.

TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen
 JOURNAL Unpublished (2001)
 COMMENT Contact: Doil Choi

Genome Research Center and National Center for Genome Information
 Korea Research Institute of Bioscience and Biotechnology
 P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
 Tel: 82-42-860-4340
 Fax: 82-42-860-4309
 Email: doil@mail.kribb.re.kr
 Plate: 068 row: E column: 11.
 Location/Qualifiers
 1..298
 /organism="Capsicum annum"
 /mol_type="mRNA"
 /db_xref="taxon:4072"
 /clone_lib="KS09"

FEATURES
 source

1..298
 /organism="Capsicum annum"
 /mol_type="mRNA"
 /db_xref="taxon:4072"
 /clone_lib="KS09"

ORIGIN

Alignment Scores:
 Pred. No.: 2.52e+03 Length: 298

Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x CAS16999 (1-298)

QY 1 GlyTyr***ValGluGlu 6
||||| |||||||

Db 235 GGATATTCTGTAGAGAA 252
||||| |||||||

RESULT 47
BU017825 300 bp mRNA linear EST 23-AUG-2002

LOCUS QHE16P11.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
DEFINITION clone QHE16P11, mRNA sequence.

ACCESSION BU017825
VERSION BU017825
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)

ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.

REFERENCE
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
1 (bases 1 to 300)
Leeds/Wash U Moss EST Project
Contact: Ralph Quatrano
Leeds/Wash U Moss EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

TITLE
JOURNAL
COMMENT

FEATURES
source
1. .300
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHE16P11"
/lab_host="E.Coli"
/clone_lib="QH_EFGHJ sunflower RHA280"
/note="Vector: pRcDNASf1AB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/TAG_TISSUE-flowers_pre-fertilized
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=GCTTGACGGG"

ORIGIN
Alignment Scores:
Pred. No.: 2.54e+03 Length: 300
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0

DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BU017825 (1-300)

QY 1 GlyTyr***ValGluGlu 6
||||| |||||||

Db 1 GGGTACAGTGTGGAGAG 18
||||| |||||||

RESULT 48
AW098038 303 bp mRNA linear EST 22-OCT-1999

LOCUS AW098038
DEFINITION AW04e02.v1 Moss EST library CPU Ceratodon purpureus cDNA clone
PEP SOURCE ID:CPU011503 5' similar to TR:082061 082061 PROTEIN
INVOLVED IN STARCH METABOLISM PRECURSOR. ;, mRNA sequence.

ACCESSION AW098038
VERSION AW098038.1 GI:6068298
KEYWORDS EST.
SOURCE Ceratodon purpureus

ORGANISM
Ceratodon purpureus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Dicranidae; Dicranales; Ditrichaceae; Ceratodon.
1 (bases 1 to 303)

REFERENCE
AUTHORS Quatrano,R., Bashardes,S., Cove,D., Cumig,A., Knight,C.,
Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T.,
Underwood,K., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Steptoe,M., Gibbons,M., Harvey,N., Ritter,E.,
Jackson,Y., McCann,R., Waterston,R. and Wilson,R.
Leeds/Wash U Moss EST Project
Other ESTs: ga04e02.x1
Unpublished (1999)
Contact: Ralph Quatrano
Leeds/Wash U Moss EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

TITLE
JOURNAL
COMMENT

FEATURES
source
1. .303
Location/Qualifiers
/organism="Ceratodon purpureus"
/mol_type="mRNA"
/db_xref="taxon:3225"
/clone="PEP_SOURCE_ID:CPU011503"
/tissue_type="protonemata: 7 day old tissue"
/lab_host="DH10B"
/clone_lib="Moss EST library CPU"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Construction of the cDNA library was carried out using Stratagene's UniZAP - cDNA synthesis kit. cDNA was constructed using an oligo dt primer/linker that contains a XhoI site within it. Following ds cDNA synthesis, EcoRI adapters were ligated to the blunt ends and sample was digested with XhoI. The result is cDNA with an EcoRI sticky end on one side and a XhoI sticky end on the other. This cDNA was ligated directionally in UniZAP arms. The vector is designed containing the pBluescript sequence as well as lambda DNA and cDNA is cloned within this pBluescript sequence. The vector was then packaged using Gold gigapackaging extracts. Library was grown in XL1blue MRF' cells and amplified. The library was excised by mass excision using Stratagene's Mass excision kit' that uses exsist as a helper phage that releases the pBluescript sequence and circularises it as single stranded plasmids that are then packaged (by helper phage) and secreted out of the host cell as phagemids. SOLR cells were transformed with phagemids and the library was plated out on LB-amp plates to select for transformants. Approximately

1,000,000 colonies were grown and recovered. The double stranded plasmid library was recovered by using Quiagen Midi prep kit. 2 micro grams of each library were used to transform DH10B cells by electroporation."

ORIGIN

Alignment Scores:
 Pred. No.: 2.57e+03 Length: 303
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 2 Gaps: 0

US-10-030-194A-5 (1-6) x AW098038 (1-303)

Qy 1 GlyTyr***ValGluGlu 6
 ||||| |||||
 Db 86 GGATACGCTGTCGAGGAG 103

RESULT 49 BB5311239 303 bp mRNA linear EST 29-JUL-2000
 LOCUS BB531239 RIKEN full-length enriched, 0 day neonate lung Mus
 DEFINITION musculus cDNA clone E030015J02 3' similar to X52101 Mouse mRNA for a 25kDa nuclear protein found in murine plasmacytoma, mRNA sequence.

ACCESSION BB531239
 VERSION BB531239.1 GI:9583168

KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugabara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Konno,H., et al.)

TITLE

JOURNAL

COMMENT

Unpublished (2000)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoaka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Thermotabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh,M., Kusunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers
 1..303
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="E030015J02"
 /tissue_type="lung"
 /dev_stage="0 day neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 0 day neonate lung"
 /note="Site 1: Sall; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
 GAGAGAGAGCGCGCAACTCGAGTTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'
 GAGAGAGAGATTCTCGAGTTAATAATATCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Alignment Scores:
 Pred. No.: 2.57e+03 Length: 303
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 2 Gaps: 0

US-10-030-194A-5 (1-6) x BB5311239 (1-303)

Qy 1 GlyTyr***ValGluGlu 6
 ||||| |||||
 Db 42 GGATACCTCTGTCGAGGAG 25

RESULT 50

LOCUS

CB700488 303 bp mRNA linear EST 10-APR-2003
 DEFINITION AMGNNUC:NRHVS-00288-H2-A W Rat hypothalamus (10471) Rattus norvegicus cDNA clone nrhys-00288-h2 5', mRNA sequence.

ACCESSION CB700488.1 GI:29757636

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Angen EST Program.
 Angen Rat EST Program
 Unpublished (2003)
 Contact: Dan Fitzpatrick
 Angen, Inc
 One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4881
 Plate: 00288 row: h column: 2.

FEATURES

source

Location/Qualifiers
 1..303
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone_lib="W Rat hypothalamus (10471)"
 /note="vector: pSPORT1; Site 1: Sall; Site 2: NotI; W Rat hypothalamus adult female Wistar rat avg. insert size 2.3

kb fraction 6 and 7"

```

ORIGIN
Alignment Scores:      2.57e+03      Length:      303
Pred. No.:            27.00      Matches:      5
Score:                83.33%      Conservative: 0
Percent Similarity:   83.33%      Conservat:    1
Best Local Similarity: 83.33%      Mismatches:  0
Query Match:         96.43%      Indels:      0
DB:                  6           Gaps:         0

US-10-030-194A-5 (1-6) x CB700488 (1-303)

Qy 1 GlyTyr***ValGluGlu 6
    ||||| ||||| ||||| |||||
Db 68 GGTACGCGAGTTGAAGAG 85

RESULT 51
CO415454/C
LOCUS Mdf3006d24.y3 Mdf3006d24 y3 mRNA linear EST 02-JUL-2004
DEFINITION Mdf3006d24.y3 Mdf3006d24 y3 Mdf3006d24 y3, mRNA sequence.
ACCESSION CO415454
VERSION CO415454.1 GI:49631702
KEYWORDS EST.
SOURCE Malus x domestica (cultivated apple)
ORGANISM Malus x domestica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 304)
Korban,S., Vodkin,L., Liu,L., Gasic,K., Gonzales,O., Hernandez,A.,
Alwinckle,H., Malnoy,M., Carroll,N., Goldsbrough,P., Orvis,K.,
Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T.,
Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Ronko,I.,
Tsgareishvili,R., Kennedy,S., Waterston,R. and Wilson,R.
Apple Functional Genomics Grant - NSF 0321702
Unpublished (2004)
TITLE Apple Functional Genomics Grant - NSF 0321702
JOURNAL
COMMENT Contact: Schuyler S. Korban
Apple Functional Genomics Grant - NSF 0321702
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library materials provided by: Schuyler S. Korban Library
constructed by: D.O.Gonzalez/L. Vodkin Library sequenced by:
Washington University Genome Sequencing Center
WASHU EST name: aah90b12.y3
Seq primer: -40RP from Gibco
High quality sequence stop: 304.
Location/Qualifiers
1..304
/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="Mdf3006d24"
/lab_host="DH10B ampicillin resistant"
/clone_lib="Mdf3"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; Total
RNA was extracted from freeze dried fruit tissue Stage 1-
Young Fruitlet < 1 cm, using the method described by Wang
and Vodkin (Plant Molecular Biology Reporter 12:132-145,
1994). Poly(A)+mRNA was isolated from total RNA using the
PolyAtract mRNA Isolation system III (Promega). The
library was prepared using the Invitrogen Life
Technologies, Superscript Plasmid System with Gateway
technology for cDNA Synthesis and cloning. Complementary
DNA was synthesized from mRNA using a Poly (dT) sequence
with a NotI restriction site. Sal I linker adapters were
ligated to the blunt ended cDNA fragments followed by
restriction with NotI. The cDNA fragments were
directionally cloned into the NotI-SalI restriction site

```

of the pSPORT 1 vector. The ligated cDNA fragments were transformed into E. coli ElectroMax DH10B host cells. Transformation efficiency: 10E+06 colony forming units. Average insert size by PCR: 992 bp"

```

ORIGIN
Alignment Scores:      2.58e+03      Length:      304
Pred. No.:            27.00      Matches:      5
Score:                83.33%      Conservative: 0
Percent Similarity:   83.33%      Conservat:    1
Best Local Similarity: 83.33%      Mismatches:  0
Query Match:         96.43%      Indels:      0
DB:                  7           Gaps:         0

US-10-030-194A-5 (1-6) x CO415454 (1-304)

Qy 1 GlyTyr***ValGluGlu 6
    ||||| ||||| ||||| |||||
Db 250 GGTATGCTGTGAGAGAG 233

RESULT 52
CR469122/C
LOCUS CR469122 Rat pBluescript Lion Rattus norvegicus cDNA clone
DEFINITION CR469122 Rat pBluescript Lion Rattus norvegicus cDNA clone
ACCESSION CR469122.1 GI:49601471
VERSION CR469122.1
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 308)
Henrich,J., Hermanns,J., Kranz,H., Loebbert,R., Schlueter,T.,
Schuette,D., Weindel,M., Heil,O., Ebert,L., Neubert,P., Peters,M.,
Radelof,U., Schneider,D. and Korn,B.
Rat ArrayTAG cDNA
Unpublished (2004)
CONTACT: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; LIONp463E08209.
RZPDLIB;
Rat ArrayTAG cDNA
http://www.rzpd.de/cgi-
bin/products/showlib.pl.cgi/response?libNo=463 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
RP: CAGGAACAGCTATGAC.
FEATURES
source
1..308
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="LIONp463E08209"
/lab_host="DH10B"
/clone_lib="Rat pBluescript Lion"

ORIGIN
Alignment Scores:      2.62e+03      Length:      308
Pred. No.:            27.00      Matches:      5
Score:                83.33%      Conservative: 0
Percent Similarity:   83.33%      Conservat:    1
Best Local Similarity: 83.33%      Mismatches:  0
Query Match:         96.43%      Indels:      0
DB:                  7           Gaps:         0

```

```

US-10-030-194A-5 (1-6) x CR469122 (1-308)

Qy 1 GlyTyr***ValGluGlu 6
Db 232 GGCTACTCTGTGAAGAG 215

RESULT 53
LOCUS AQ233114/c
DEFINITION HS_2037_A1_F08_MR_CIT Approved Human Genomic Sperm Library D Homo
          sapiens_genomic clone Plate=2037 Col=15 Row=K, genomic survey
          sequence.
ACCESSION AQ233114
VERSION AQ233114.1 GI:3658343
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 309)
          Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and
          Keller,A., Shaker,J., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
          Hood,L.
          Sequence-tagged connectors: A sequence approach to mapping and
          scanning the human genome
          Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
          99380589
          10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
          High Throughput Sequencing Center
          University of Washington
          401 Queen Anne Avenue North, Seattle, WA 98109, USA
          Tel: (206) 616-3618
          Fax: (206) 616-3887
          Email: jwallace@u.washington.edu
          Sequence Tagged Connector
          Plate: 2037 row: K column: 15
          Class: BAC ends
          High quality sequence stop: 309.

FEATURES             source
    source
    1..309
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /clones="Plate=2037 Col=15 Row=K"
        /sex="male"
        /clone_lib="CIT Approved Human Genomic Sperm Library D"
        /note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
          E-Coli DH10B"

ORIGIN
Alignment Scores:
Pred. No.:          2.63e+03      Length:          309
Score:              27.00         Matches:          5
Percent Similarity: 83.33%        Conservative:    0
Best Local Similarity: 83.33%      Mismatches:       1
Query Match:        96.43%         Gaps:            0
DB:
Gaps:              8
Gaps:              0

US-10-030-194A-5 (1-6) x AQ233114 (1-309)

Qy 1 GlyTyr***ValGluGlu 6
Db 157 GGTATGCAGTAGAGAA 140

RESULT 54
LOCUS A1715497/c
DEFINITION A1715497 317 bp mRNA linear EST 10-JUN-1999
          UI-R-Y0-abq-c-06-0-UI.s1 UI-R-Y0 Rattus norvegicus cDNA clone
          sequence.
ACCESSION A1715497
VERSION A1715497.1 GI:5032750

```

KEYWORDS SOURCE ORGANISM

EST.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE AUTHORS

1 (bases 1 to 317)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
MEDLINE
PUBMED
COMMENT

JOURNAL MEDLINE PUBMED COMMENT

Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
normalized Eye library cDNA Library Preparation: M.B. Soares Lab
Clone distribution: clones will be available through Research
Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES source

Location/Qualifiers
1..317
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-Y0-abq-c-06-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-Y0"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-Y0
library is a subtracted library derived from an
individually-tagged normalized whole-eye (minus the lens)
library. The driver for the subtraction consisted of a
pool of all previous libraries (UI-R-A0, UI-R-A1, UI-R-E0,
UI-R-E1, UI-R-C0, and UI-R-C1). The tag is a string of
3-5 nucleotides present between the Not I site and the
oligo-dT track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-Y0) was constructed as follows: PCR
amplified cDNA inserts from previous library clones from
which 3' ESTs had been derived were used as a driver in a
hybridization with the normalized whole-eye library in
the form of single-stranded circles. The remaining
single-stranded circles (subtracted library) was purified
by hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (life technologies) to generate the UI-R-Y0
library. This procedure has been previously described
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
1996)
TAG TISSUE=Eye
TAG_LIB=UI-R-Y0
TAG_SEQ=CAITG"

ORIGIN

Alignment Scores:
Pred. No.: 2.71e+03 Length: 317
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0

```

DB:
US-10-030-194A-5 (1-6) x AI715497 (1-317)

Qy 1 GlyTyr***ValGluGlu 6
    |||||
Db 239 GGCTACTCTGTGAAGAG 222

RESULT 55
CF034326
LOCUS QCF3h10.yg QCF Zea mays cDNA clone QCF3h10, mRNA linear EST 17-JUL-2003
DEFINITION QCF3h10.yg QCF Zea mays cDNA clone QCF3h10, mRNA sequence.
ACCESSION CF034326
VERSION CF034326.1 GI:32929514
KEYWORDS EST.
SOURCE
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 317)
REFERENCE
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com)
and http://genoplante-info.infobiogen.fr.
FEATURES
source
Location/Qualifiers
1..317
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="P2"
/db_xref="taxon:4577"
/clone="QCF3h10"
/tissue_type="seedling minus kernel"
/clone_lib="QCF"

ORIGIN
Alignment Scores:
Pred. No.: 2,71e+03 Length: 317
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-5 (1-6) x CF034326 (1-317)

Qy 1 GlyTyr***ValGluGlu 6
    |||||
Db 20 GGCTACTCCGTGAAGAA 37

RESULT 56
BQ146842
LOCUS BQ146842/2 320 bp mRNA linear EST 24-APR-2002
DEFINITION NF029A03FL1020 Developing flower Medicago truncatula cDNA clone
ACCESSION NF029A03FL 5', mRNA sequence.
BQ146842
VERSION BQ146842.1 GI:20283901
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 320)
REFERENCE
AUTHORS Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,

```

```

Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula flower library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 320 Std Error: 0.00
Plate: 029 Row: A Column: 03
Seq primer: TCACACGAGAAACAGCTATGAC.
FEATURES
source
Location/Qualifiers
1..320
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF029A03FL"
/tissue_type="Developing flowers"
/dev_stage="Developmentally pooled. Contains a mixture of
very young, developing, fully-opened flowers and flowers
in early transition into pods."
/clone_lib="Developing flower"
/note="Vector: Lambda Zap; cDNA was prepared from polyA+
enriched, pooled samples of equivalent amounts of total
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the Gigapack III Gold
packaging extracts. Phagemids containing cDNA inserts were
in vivo excised from the recombinant Uni-Zap XR vector
using ExAssist helper phage and the E. coli strain
XL1-Blue MRP' (Stratagene). Excised plasmids were plated
using SOLR cells."

ORIGIN
Alignment Scores:
Pred. No.: 2,74e+03 Length: 320
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BQ146842 (1-320)

Qy 1 GlyTyr***ValGluGlu 6
    |||||
Db 74 GGATATACAGTAGAAGAA 57

RESULT 57
BQ911150
LOCUS BQ911150 321 bp mRNA linear EST 19-AUG-2002
DEFINITION QHA16D20.yg.ab1 OH ABCDI sunflower RHA801 Helianthus annuus cDNA
clone QHA16D20, mRNA sequence.
BQ911150
ACCESSION BQ911150.1 GI:22309929
VERSION BQ911150.1
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
1 (bases 1 to 321)
REFERENCE
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Riesberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/

```



```

/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clones="QHA13P08"
/lab_host="E.coli"
/clone_lib="QH_ABCDI sunflower RHA801"
/note="vector: pBRCNASTAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=chemical induction
TAG_LIB=QH ABCDI sunflower RHA801
TAG_SEQ-TGTAGCGGG"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2.78e+03 Length: 324
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

```

US-10-030-194A-5 (1-6) x BQ910364 (1-324)

```

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 1 GGGTACAGTGTGAAGAG 18

```

```

RESULT 60
CD476039/c
LOCUS
DEFINITION CD476039 333 bp mRNA linear EST 04-JUN-2003
nad03-2msl-e05 Nad03 Nuphar advena cDNA clone nad03-2msl-e05 5',
mRNA sequence.
CD476039
CD476039.1 GI:31397307
EST.
Nuphar advena
Nuphar advena

```

```

ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; stem Magnoliophyta; Nymphaeales;
Nymphaeaceae; Nuphar.
1 (bases 1 to 333)
dePamphilis,C., Carlson,J., Ma,H., Frohlich,M., Tankeley,S.,
Leebens-Mack,J., Field,D., Arrington,J., Zahn,L., Kong,H.,
Druckemiller,M., Landherr,L., Hu,Y., Ilut,D., Wall,K.,
Plock,S.Chioean,S.,Albert,V., Doyle,J., Miller,W.,
Oppenheimer,D., Soltis,D., Soltis,P. and Theissen,G.
Generation of ESTs from early flower buds of Nuphar advena
Unpublished (2002)
Contact: Claude dePamphilis or James Leebens-Mack
Mueller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 6413
Fax: 814 865 9131
Email: cwd3@psu.edu or jhl10@psu.edu

```

```

The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (http://pgn.cornell.edu)
Plate: nad03-2msl row: e column: 05
Seq primer: M13F.

```

FEATURES

```

source
1..333
/organism="Nuphar advena"
/mol_type="mRNA"

```

```

/db_xref="taxon:77108"
/clone="nad03-2msl-e05"
/tissue_type="flower buds"
/dev_stage="c= 2.5mm"
/lab_host="SOLR"
/clone_lib="Nad03"
/note="vector: pBluescript SK (+/-); Site 1: EcoRI;
Site 2: XhoI; Only floral buds with diameter of 2.5 mm of
less were used for RNA isolation. This is a directionally
cloned, non-normalized library. Avg. insert length: 1134;
Primers: M13F and M13R; Antibiotic: 50 ug/ml Ampicillin;
Primary Titer: 256 pfu total; Amplified Titer: 3.2810
pfu/ml; Mass Excised Titer: 5810 total; This library has
been generated by the Floral Genome Project (FGP). The
Floral Genome Project is funded by NSF's Plant Genome
Research Program (DBI-0115684). More information about the
project can be obtained at http://fgp.bio.psu.edu"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2.88e+03 Length: 333
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 6 Gaps: 0

```

US-10-030-194A-5 (1-6) x CD476039 (1-333)

```

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 222 GGTACAGTGTGAAGAG 205

```

RESULT 61

```

R31872 333 bp mRNA linear EST 28-APR-1995
LOCUS
DEFINITION Y595e10.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
IMAGE134058 5', mRNA sequence.
R31872
R31872.1 GI:787715
EST.
Homo sapiens (human)

```

```

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 333)
Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1095
High quality sequence stops: 203
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1095 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 203.
Location/Qualifiers
1..333
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:539842"
/db_xref="taxon:9606"

```

TITLE

JOURNAL

COMMENT

/clones="IMAGE:134058"
 /sex="female"
 /dev stages="placenta obtained at birth (full term)"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_libs="Soares placenta NB2HP"
 /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGATTCCGCGCCGACGAAATTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo. "

ORIGIN

Alignment Scores:
 Pred. No.: 2.88e+03 Length: 333
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservatave: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 7 Gaps: 0

US-10-030-194A-5 (1-6) x R31872 (1-333)

Qy 1 GlyTyr***ValGluGlu 6
 ||||| |||||
 Db 281 GGTACACAGTTGAGGAG 298

RESULT 62

BQ974743

LOCUS

DEFINITION BQ974743 335 bp mRNA linear EST 21-AUG-2002
 clone QH16124, sunflower RHA801 Helianthus annuus cDNA

ACCESSION BQ974743

VERSION BQ974743.1

KEYWORDS GI:22392266

SOURCE EST.

ORGANISM Helianthus annuus (common sunflower)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.

REFERENCE 1 (bases 1 to 335)

AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.

TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project

http://compgenomics.ucdavis.edu/

Unpublished (2002)

COMMENT Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)

Asmundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@atgc.org [michelmore@vegmil.ucdavis.edu]

belongs to contig QH_CA_Contig2725, see http://cgdb.ucdavis.edu/ for details.

Plate: QH16 row: I column: 24.

FEATURES

source

1. .335
 Location/Qualifiers
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA801"
 /db_xref="taxon:4232"
 /clones="QH16124"
 /lab_host="E.coli"
 /clone_libs="QH ABCDI sunflower RHA801"
 /note="vector: pBRCDNASFIAB; The library was constructed from 11 different sources of RNA from a single genotype.

ORIGIN

Alignment Scores:

Pred. No.: 2.9e+03 Length: 335
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservatave: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BQ974743 (1-335)

Qy 1 GlyTyr***ValGluGlu 6
 ||||| |||||
 Db 55 GGTACACAGTTGGAAGAG 72

RESULT 63

BU016795

LOCUS

DEFINITION BU016795 337 bp mRNA linear EST 23-AUG-2002
 clone QHE14A15, mRNA sequence.

ACCESSION BU016795

VERSION BU016795.1

KEYWORDS GI:22452315

SOURCE EST.

ORGANISM Helianthus annuus (common sunflower)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.

REFERENCE 1 (bases 1 to 337)

AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.

TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project

http://compgenomics.ucdavis.edu/

Unpublished (2002)

COMMENT Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)

Asmundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@atgc.org [michelmore@vegmil.ucdavis.edu]

belongs to contig QH_CA_Contig2725, see http://cgdb.ucdavis.edu/ for details.

Plate: QHE14 row: A column: 15.

FEATURES

source

1. .337
 Location/Qualifiers
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA280"
 /db_xref="taxon:4232"
 /clones="QHE14A15"
 /lab_host="E.coli"
 /clone_libs="QH_EFGHJ sunflower RHA280"
 /note="vector: pBRCDNASFIAB; The library was constructed from 11 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library

Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/
 TAG_SEQ=Not found"

construction can be obtained at <http://cgdb.ucdavis.edu/>
 TAG TISSUE=shoots environmental stress
 TAG_LIB=QH_EFGUJ sunflower RHA280
 TAG_SEQ=TCGCAACGGG"

ORIGIN

Alignment Scores:
 Pred. No.: 2.92e+03 Length: 337
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservativeness: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BU016795 (1-337)

Qy 1 GlyTyr***ValGluGlu 6

Db 96 GGGTACAGTGTGGAAGAG 113

RESULT 64

BF989730 338 bp mRNA linear EST 23-JAN-2001
 LOCUS IL5-GN0178-231000-195-h09 GN0178 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF989730
 ACCESSION BF989730
 VERSION BF989730.1 GI:12396055

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 338)

REFERENCE

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL5&t2=IL5-GN0178-231000-195-h09&t3=2000-10-23&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence stop: 326.
 Location/Qualifiers

FEATURES

source

1..338
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="GN0178"
 /note="Organ: placenta normal; Vector: puc18; Site 1:
 SmaI; Site 2: SmaI; A mini-library was made by cloning
 products derived from ORESTES PCR (U.S. Letters Patent
 application No. 196,716 - Ludwig Institute for Cancer
 Research) profiles into the puc 18 vector. Reverse
 transcription of tissue mRNA and cDNA amplification were
 performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.: 2.93e+03 Length: 338
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservativeness: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x BF989730 (1-338)

Qy 1 GlyTyr***ValGluGlu 6

Db 247 GGATATGCAGTAGAGAA 264

RESULT 65

LOCUS

BF989729 339 bp mRNA linear EST 23-JAN-2001
 IL5-GN0178-231000-195-h08 GN0178 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF989729

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT

LABORATORY

LUDWIG INSTITUTE

RUA PROF. ANTONIO PRUDENTE

BRAZIL

TEL:

FAX:

EMAIL:

THIS SEQUENCE

PROJECT.

THIS ENTRY

URL

231000-195-H08&T3=2000-10-23&T4=1

SEQ PRIMER:

HIGH QUALITY

FEATURES

source

1..339

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="GN0178"

/note="Organ: placenta normal; Vector: puc18; Site 1:

SmaI; Site 2: SmaI; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the puc 18 vector. Reverse

transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

Length:

Matches:

Conservativeness:

Mismatches:

Indels:

```

DB:
US-10-030-194A-5 (1-6) x BF989729 (1-339)

Qy 1 GlyTyr***ValGluGlu 6
Db 247 GGTATGCAGTAGAGAA 264

RESULT 66
BU017766
LOCUS
DEFINITION
clone QHE16M20.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
clone QHE16M20, mRNA sequence.
ACCESSION
BU017766
VERSION
BU017766.1 GI:22453286
KEYWORDS
EST.
SOURCE
Helianthus annuus (common sunflower)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE
1 (Bases 1 to 339)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
JOURNAL
Unpublished (2002)
COMMENT
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QH_CA_Contig2725, see http://cgpdb.ucdavis.edu/
for details.
Plate: QHE16 row: M column: 20.
FEATURES
source
1..339
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHE16M20"
/lab_host="E.coli"
/clone_lib="QH_EFGHJ sunflower RHA280"
/note="Vector: pBRCNASTiAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=shoots environmental stress
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCAACGGG"

ORIGIN
Alignment Scores:
Pred. No.: 2,94e+03 Length: 339
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BU017766 (1-339)

```

```

Qy 1 GlyTyr***ValGluGlu 6
Db 96 GGTACAGTGTGGAAGAG 113

RESULT 67
BQ975877
LOCUS
DEFINITION
OH1M14.YG.ab1 QH_ABCDI sunflower RHA801 Helianthus annuus cDNA
clone OH1M14, mRNA sequence.
ACCESSION
BQ975877
VERSION
BQ975877.1 GI:22393400
KEYWORDS
EST.
SOURCE
Helianthus annuus (common sunflower)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE
1 (Bases 1 to 341)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
JOURNAL
Unpublished (2002)
COMMENT
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QH_CA_Contig2725, see http://cgpdb.ucdavis.edu/
for details.
Plate: QH11 row: M column: 14.
FEATURES
source
1..341
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="QH1M14"
/lab_host="E.coli"
/clone_lib="QH_ABCDI sunflower RHA801"
/note="Vector: pBRCNASTiAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=huills
TAG_LIB=QH_ABCDI sunflower RHA801
TAG_SEQ=GCTAGTCGGG"

ORIGIN
Alignment Scores:
Pred. No.: 2,96e+03 Length: 341
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BQ975877 (1-341)

Qy 1 GlyTyr***ValGluGlu 6
Db 1 GGTACAGTGTGGAAGAG 18

```

```

RESULT 68
BU024708
LOCUS
DEFINITION QHF5L21.Yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus CDNA
ACCESSION BU024708 349 bp mRNA linear EST 23-AUG-2002
SOURCE clone QHF5L21, mRNA sequence.
ORGANISM Helianthus annuus (common sunflower)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
1 (bases 1 to 349)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compenomics.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QH_CA_Contig2725, see http://cgdb.ucdavis.edu/
for details.
Plate: QHF5 row: L column: 21.
FEATURES
source
1..349
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHF5L21"
/lab_host="E.coli"
/clone_lib="QH_EFGHJ sunflower RHA280"
/note="Vector: pRcDNASfiAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgdb.ucdavis.edu/
TAG_TISSUE=flowers pre-fertilized
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=GCTTGACGGG"

ORIGIN
Alignment Scores:
Pred. No.: 3.04e+03 Length: 349
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BU024708 (1-349)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||
Db 1 GGGTACAGTGTGGAAGAG 18

RESULT 69
BU025543
LOCUS
DEFINITION QHF9P05.Yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus CDNA
ACCESSION BU025543 351 bp mRNA linear EST 26-JUN-2002
SOURCE faa94a09.y1 Sugano SJD adult male Danio rerio cDNA clone
IMAGE:5916208 5' similar to TR:Q99627 Q99627 HCOP9. i, mRNA
sequence.
BQ616698
BQ616698.1 GI:21606367

```

```

ACCESSION BU025543 mRNA sequence.
VERSION BU025543.1 GI:22461063
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
1 (bases 1 to 350)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compenomics.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QH_CA_Contig2725, see http://cgdb.ucdavis.edu/
for details.
Plate: QHF9 row: P column: 05.
FEATURES
source
1..350
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHF9P05"
/lab_host="E.coli"
/clone_lib="QH_EFGHJ sunflower RHA280"
/note="Vector: pRcDNASfiAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgdb.ucdavis.edu/
TAG_TISSUE=flowers pre-fertilized
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=GCTTGACGGG"

ORIGIN
Alignment Scores:
Pred. No.: 3.05e+03 Length: 350
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-5 (1-6) x BU025543 (1-350)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||
Db 1 GGGTACAGTGTGGAAGAG 18

RESULT 70
BQ616698
LOCUS
DEFINITION BQ616698 351 bp mRNA linear EST 26-JUN-2002
IMAGE:5916208 5' similar to TR:Q99627 Q99627 HCOP9. i, mRNA
sequence.
BQ616698
BQ616698.1 GI:21606367

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

EST.
 Danio rerio (zebrafish)
 Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 351)
 Clark M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M.,
 Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y.,
 Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
 Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R. and Wilson, R.
 WashU Zebrafish EST Project 1998
 Unpublished (1998)
 Other ESTs: faa94a09.xl
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrafish@wustl.wustl.edu
 Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
 Sequencing by: Washington University Genome Sequencing Center Clone
 distribution information can be found through the I.M.A.G.E.
 Consortium/LiNL, send email to: info@image.llnl.gov
 Seq primer: T3 ET from Amersham
 High quality sequence stop: 192.

FEATURES
 source
 1. .351
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:5916208"
 /sex="male"
 /tissue_type="whole body"
 /dev_stage="adult"
 /lab_host="DH10B (phage resistant)"
 /clone_lib="Sugano SJD adult male"
 /note="Vector: pME18S-FL3; Site 1: DraIII (CACCATGTG);
 Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed
 with an oligo(dT) primer (ATGTGGCTTTTCTTTTCTTTT);
 double-stranded cDNA was ligated to a DraIII adaptor
 [TGTGTGCGTACTG], digested and cloned into distinct DraIII
 sites of the pME18S-FL3 vector (5' site CACCATGTG, 3' site
 CACCATGTG). XhoI should be used to isolate the cDNA
 insert. Size selection was performed to exclude fragments
 <1.5kb. Library constructed and donated by Dr. Sumio
 Sugano (University of Tokyo Institute of Medical Science).
 Custom primers for sequencing: 5' end primer
 CTTCTGCTCTTAAGCTGCG and 3' end primer
 CGACCTGAGCTCGACACA."

ORIGIN
 Alignment Scores:
 Pred. No.: 3.06e+03 Length: 351
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 5 Gaps: 0
 US-10-030-194A-5 (1-6) x BQ616698 (1-351)
 Qy 1 GlyTyr***ValGluGlu 6
 |||||
 Db 117 GGATACACTGGTGAAGAG 134
 RESULT 71
 BU018981
 LOCUS
 DEFINITION
 BU018981
 QHE1907.YG.ab1 QH EFGHJ sunflower RHA280 Helianthus annuus cDNA
 clone QHE1907, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BU018981
 BU018981.1 GI:22454501
 EST.
 Helianthus annuus (common sunflower)
 Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Helianthus.
 1 (bases 1 to 351)
 Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
 Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
 Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
 Lai, Z., Church, S., Jackson, L. and Bradford, K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compgenomics.ucdavis.edu/
 Unpublished (2002)
 Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QH_CA_Contig2725, see http://cgpdb.ucdavis.edu/
 for details.
 Plate: QHE1 row: g column: 07.

FEATURES
 source
 1. .351
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA280"
 /db_xref="taxon:4232"
 /clone="QHE1907"
 /lab_host="E.coli"
 /clone_lib="QH EFGHJ sunflower RHA280"
 /note="Vector: pBRCNDSfiAB; The library was constructed
 from 11 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each
 source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and
 transformations made with four size classes to minimize
 size bias. Details of each source of RNA and library
 construction can be obtained at http://cgpdb.ucdavis.edu/
 TAG_TISSUE=chemical induction
 TAG_LIB=QH EFGHJ sunflower RHA280
 TAG_SEQ=TTGTAGCCGGG"

ORIGIN
 Alignment Scores:
 Pred. No.: 3.06e+03 Length: 351
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 5 Gaps: 0
 US-10-030-194A-5 (1-6) x BU018981 (1-351)
 Qy 1 GlyTyr***ValGluGlu 6
 |||||
 Db 1 GGTACACTGTGGAAGAG 18
 RESULT 72
 AW294420
 LOCUS
 DEFINITION
 AW294420
 IMAGE:2726606 3', mRNA sequence.
 ACCSSION
 AW294420
 VERSION
 AW294420.1 GI:6701056
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 353)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb-remail.nih.gov
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward
 POLYA-Yes.

FEATURES
 source
 1..353
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2726606"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Sub4"
 /note="Vector: p7T73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; The NCI CGAP Sub4 library is a subtracted library derived from the NCI CGAP Sub2 library which is a subtracted library derived from the NCI CGAP Sub1 library, which is a subtracted library derived from B1. B1 constitutes a mixture of 21 normalized or subtracted NCI CGAP libraries: NCI CGAP Co4, NCI CGAP Pr22, NCI CGAP Pr28, NCI CGAP Co10, NCI CGAP Co16, NCI CGAP Kid5, NCI CGAP Kid12, NCI CGAP Kid3, NCI CGAP Kid11, NCI CGAP Lym2, NCI CGAP Br2, NCI CGAP Co8, NCI CGAP CLL1, NCI CGAP Le12, NCI CGAP Brn23, NCI CGAP Lu5, NCI CGAP Lu24, NCI CGAP Lu19, NCI CGAP GC4, NCI CGAP GC6, NCI CGAP Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
 NCI CGAP Kid3 pool 1 : LLAM 3334-3337, 3682-3683,
 3798-3803 (IMAGE CloneIDs 1322376-1323911,
 1456008-1456775, 1500552-1502855) NCI CGAP Kid5 pool 1 :
 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs 132912-1325831, 1471368-1472903, 1492104-1493255)
 NCI CGAP Lu5 pool 1 : LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439) NCI CGAP GC4 pool 1 : LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631, 1469064-1470983,
 1475592-1476743) NCI CGAP Pr22 pool 1 : LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759,
 1101192-1101959, 1217928-1220615) NCI CGAP Co10 pool 1 :
 LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255,
 1144584-1145351) Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6, 791-806.]
 TAG_TISSUE=lung
 TAG_LIB=NCI CGAP_Lu5
 TAG_SEQ=CACAC"

ORIGIN
 Alignment Scores:
 Pred. No.: 3.08e+03 Length: 353
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 2 Gaps: 0

US-10-030-194a-5 (1-6) x AW294420 (1-353)
 Qy 1 GlyTyr***ValGluGlu 6
 ||||| |||||
 Db 287 GGTATTCTGTGTGAAGAA 304

RESULT 73
 BQ910337 355 bp mRNA linear EST 19-AUG-2002
 LOCUS QHA13002.Yg.ab1 QH ABCDI sunflower RHA801 Helianthus annuus cDNA
 DEFINITION clone QHA13002, mRNA sequence.
 ACCSSION BQ910337
 VERSION BQ910337.1 GI:22309116
 KEYWORDS EST.
 SOURCE Helianthus annuus (common sunflower)
 ORGANISM Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.
 REFERENCE 1 (bases 1 to 355)
 AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.
 TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
 JOURNAL Unpublished (2002)
 COMMENT Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QH_CA_Contig2725, see <http://cgdb.ucdavis.edu/> for details.
 Plate: QHA13 row: 0 column: 02.
 Location/Qualifiers
 1..355
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA801"
 /db_xref="taxon:4232"
 /clone="QHA13002"
 /lab_host="E.coli"
 /clone_lib="QH ABCDI sunflower RHA801"
 /note="Vector: pBRCNASTAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgdb.ucdavis.edu/>
 TAG_TISSUE=chemical induction
 TAG_LIB=QH ABCDI sunflower RHA801
 TAG_SEQ=GTAGCCGGG"

ORIGIN
 Alignment Scores:
 Pred. No.: 3.1e+03 Length: 355
 Score: 27.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.43% Indels: 0
 DB: 5 Gaps: 0

US-10-030-194a-5 (1-6) x BQ910337 (1-355)
 Qy 1 GlyTyr***ValGluGlu 6

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2004, 13:41:12 ; Search time 72.5 Seconds
(without alignments)
58.824 Million cell updates/sec

Title: US-10-030-194A-5
Perfect score: 28
Sequence: 1 GYXVEE 6

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEVS=xlp
-Q=/cgn2_1/USPTO spool_p/US10030194/runat_01112004.184848.28987/app_query.fasta_1.398
-DB=Issued Patents_NA -QFMT=fastcap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCUT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi
-LIST=1000 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10030194 @CGN 1.1.164 @runat_01112004.184848.28987 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRSADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	27	96.4	849	4 US-09-107-532A-1540	Sequence 1540, Ap
C 2	27	96.4	1050	4 US-09-252-991A-10973	Sequence 10973, A
C 3	27	96.4	1065	4 US-09-107-532A-1525	Sequence 1525, Ap
C 4	27	96.4	2289	2 US-08-907-166-9	Sequence 9, Appli
C 5	27	96.4	2289	4 US-09-391-340-9	Sequence 9, Appli
C 6	27	96.4	2547	4 US-09-252-991A-11271	Sequence 11271, A
C 7	27	96.4	2964	4 US-09-252-991A-10912	Sequence 10912, A
C 8	27	96.4	13321	4 US-08-956-171E-4	Sequence 4, Appli
C 9	27	96.4	13321	4 US-08-781-986A-4	Sequence 4, Appli
C 10	27	96.4	17000	4 US-09-679-299A-18	Sequence 18, Appli
C 11	26	92.9	39	1 US-08-313-050-7	Sequence 7, Appli
C 12	26	92.9	107	4 US-09-513-999C-36409	Sequence 36409, A

159	92.9	26	1	US-08-313-050-9	Sequence 9, Appli
170	92.9	26	4	US-09-513-999C-33426	Sequence 33426, A
239	92.9	26	4	US-09-621-976-9290	Sequence 9290, Ap
262	92.9	26	4	US-09-513-999C-17433	Sequence 17433, A
263	92.9	26	4	US-09-513-999C-3343	Sequence 3343, Ap
270	92.9	26	3	US-09-134-001C-1648	Sequence 1648, Ap
276	92.9	26	1	US-08-313-050-17	Sequence 17, Appli
305	92.9	26	3	US-08-808-982-4	Sequence 4, Appli
305	92.9	26	3	US-09-306-902A-4	Sequence 4, Appli
305	92.9	26	4	US-09-513-999C-22044	Sequence 22044, A
324	92.9	26	4	US-09-107-532A-2175	Sequence 2175, Ap
383	92.9	26	4	US-09-513-999C-35661	Sequence 35661, A
387	92.9	26	4	US-09-134-000C-165	Sequence 165, App
403	92.9	26	4	US-09-513-999C-15816	Sequence 15816, A
404	92.9	26	4	US-09-513-999C-28426	Sequence 28426, A
445	92.9	26	4	US-09-173-300-12	Sequence 12, Appli
455	92.9	26	4	US-09-621-976-12487	Sequence 12487, A
495	92.9	26	3	US-09-280-116-153	Sequence 153, App
575	92.9	26	4	US-09-208-258-91	Sequence 91, Appli
583	92.9	26	4	US-09-585-173B-19	Sequence 19, Appli
717	92.9	26	4	US-09-107-532A-2614	Sequence 2614, Ap
771	92.9	26	1	US-08-253-155A-17	Sequence 17, Appli
782	92.9	26	4	US-09-585-173B-21	Sequence 21, Appli
935	92.9	26	1	US-08-313-050-10	Sequence 10, Appli
993	92.9	26	4	US-09-134-000C-86	Sequence 86, Appli
1001	92.9	26	4	US-09-641-638-204	Sequence 204, App
1001	92.9	26	4	US-10-170-097-204	Sequence 204, App
1074	92.9	26	3	US-09-134-001C-1568	Sequence 1568, Ap
1077	92.9	26	4	US-09-710-279-17	Sequence 17, Appli
1086	92.9	26	4	US-09-173-300-14	Sequence 14, Appli
1089	92.9	26	3	US-09-134-001C-2566	Sequence 2566, Ap
1109	92.9	26	1	US-08-183-214-11	Sequence 11, Appli
1128	92.9	26	4	US-09-543-681A-3142	Sequence 3142, Ap
1138	92.9	26	4	US-09-270-767-28736	Sequence 28736, A
1145	92.9	26	3	US-08-808-103B-7	Sequence 7, Appli
1145	92.9	26	3	US-08-838-151A-19	Sequence 19, Appli
1145	92.9	26	3	US-08-838-151A-23	Sequence 23, Appli
1145	92.9	26	3	US-08-838-151A-26	Sequence 26, Appli
1145	92.9	26	3	US-08-838-151A-29	Sequence 29, Appli
1148	92.9	26	3	US-08-809-103B-1	Sequence 1, Appli
1150	92.9	26	3	US-08-809-103B-3	Sequence 3, Appli
1150	92.9	26	3	US-08-809-103B-5	Sequence 5, Appli
1200	92.9	26	4	US-09-252-991A-8468	Sequence 8468, Ap
1239	92.9	26	4	US-09-543-681A-3723	Sequence 3723, Ap
1252	92.9	26	4	US-09-673-395A-73	Sequence 73, Appli
1383	92.9	26	3	US-09-134-001C-2485	Sequence 2485, Ap
1476	92.9	26	3	US-08-872-302-1	Sequence 1, Appli
1636	92.9	26	3	US-09-117-853-7	Sequence 7, Appli
1636	92.9	26	4	US-09-911-154-7	Sequence 7, Appli
1636	92.9	26	4	US-09-911-154-7	Sequence 7, Appli
1642	92.9	26	3	US-09-117-853-5	Sequence 5, Appli
1642	92.9	26	3	US-09-117-853-9	Sequence 9, Appli
1642	92.9	26	4	US-09-911-154-5	Sequence 5, Appli
1642	92.9	26	4	US-09-911-154-9	Sequence 9, Appli
1642	92.9	26	4	US-09-911-154-9	Sequence 9, Appli
1643	92.9	26	3	US-09-117-853-3	Sequence 3, Appli
1643	92.9	26	4	US-09-911-154-3	Sequence 3, Appli
1643	92.9	26	4	US-09-911-154-3	Sequence 3, Appli
1768	92.9	26	4	US-09-485-529-13	Sequence 13, Appli
1771	92.9	26	2	US-09-976-594-28	Sequence 28, Appli
1837	92.9	26	4	US-08-750-134A-4	Sequence 4, Appli
1837	92.9	26	3	US-09-363-745-4	Sequence 4, Appli
1964	92.9	26	3	US-09-117-853-1	Sequence 1, Appli
1964	92.9	26	4	US-09-911-154-1	Sequence 1, Appli
2118	92.9	26	4	US-09-270-767-12884	Sequence 12884, A
2125	92.9	26	4	US-09-485-529-14	Sequence 14, Appli
2133	92.9	26	4	US-09-252-991A-8375	Sequence 8375, Ap
2203	92.9	26	1	US-08-076-091C-13	Sequence 13, Appli
2203	92.9	26	1	US-08-285-641-13	Sequence 13, Appli
2240	92.9	26	4	US-09-489-847-84	Sequence 84, Appli

86	26	92.9	2255	4	US-09-485-529-15	Sequence 15, Appl	159	25	89.3	6070	4	US-10-204-708-10	Sequence 10, Appl
87	26	92.9	2311	4	US-09-489-847-123	Sequence 123, App	160	25	89.3	10467	4	US-10-204-708-1	Sequence 1, Appli
88	26	92.9	2580	4	US-09-107-532A-3175	Sequence 3175, Ap	161	25	89.3	48763	4	US-09-916-204-3	Sequence 3, Appli
89	26	92.9	2610	4	US-09-543-681A-642	Sequence 642, App	162	25	89.3	48763	4	US-10-282-048-3	Sequence 3, Appli
90	26	92.9	2831	2	US-08-808-982-3	Sequence 3, Appli	c 163	25	89.3	246240	2	US-08-724-394A-20	Sequence 20, Appl
91	26	92.9	2831	3	US-09-306-902A-3	Sequence 3, Appli	c 164	25	89.3	246240	2	US-08-724-394A-21	Sequence 21, Appl
92	26	92.9	2885	3	US-08-948-705-2	Sequence 2, Appli	c 165	25	89.3	246240	2	US-08-724-394A-22	Sequence 22, Appl
93	26	92.9	2885	4	US-09-510-543-2	Sequence 2, Appli	166	25	89.3	1664976	4	US-08-916-421B-1	Sequence 1, Appli
94	26	92.9	2885	4	US-09-849-602-3	Sequence 3, Appli	167	25	89.3	1664976	4	US-09-692-570-1	Sequence 1, Appli
95	26	92.9	3001	4	US-09-539-333D-149	Sequence 149, App	168	25	89.3	1830121	4	US-09-557-884-1	Sequence 1, Appli
96	26	92.9	3024	2	US-08-836-943-1	Sequence 1, Appli	169	25	89.3	1830121	4	US-09-643-990A-1	Sequence 1, Appli
97	26	92.9	3204	4	US-09-710-279-3528	Sequence 3528, Ap	170	25	89.3	1830121	4	US-10-329-960-1	Sequence 1, Appli
98	26	92.9	3205	4	US-09-134-000C-2685	Sequence 2685, Ap	171	24	85.7	223	3	US-09-102-491-4	Sequence 4, Appli
99	26	92.9	3330	1	US-08-149-103-1	Sequence 1, Appli	c 172	24	85.7	266	4	US-09-702-705-885	Sequence 885, App
100	26	92.9	3531	1	US-08-451-883-1	Sequence 1, Appli	c 173	24	85.7	266	4	US-09-736-457-885	Sequence 885, App
101	26	92.9	3531	4	US-09-976-594-1083	Sequence 1083, Ap	c 174	24	85.7	266	4	US-09-614-124B-885	Sequence 885, App
102	26	92.9	3571	4	US-09-799-451-411	Sequence 411, App	c 175	24	85.7	266	4	US-09-671-325-885	Sequence 885, App
103	26	92.9	3604	4	US-09-710-279-4199	Sequence 4199, Ap	c 176	24	85.7	266	4	US-09-658-824-885	Sequence 885, App
104	26	92.9	3611	4	US-09-221-017B-877	Sequence 877, App	c 177	24	85.7	290	4	US-09-270-767-857	Sequence 857, App
105	26	92.9	3656	1	US-08-393-734-1	Sequence 1, Appli	178	24	85.7	290	4	US-09-270-767-16139	Sequence 16139, A
106	26	92.9	3656	3	US-08-894-489-1	Sequence 1, Appli	179	24	85.7	291	4	US-09-134-000C-2711	Sequence 2711, A
107	26	92.9	3884	4	US-10-140-002-145	Sequence 145, App	180	24	85.7	306	3	US-09-134-000C-1026	Sequence 1026, Ap
108	26	92.9	3966	4	US-09-976-594-555	Sequence 555, App	c 181	24	85.7	342	3	US-08-929-856-188	Sequence 65, Appl
109	26	92.9	4090	4	US-09-710-279-3634	Sequence 3634, Ap	c 182	24	85.7	342	3	US-08-929-856-188	Sequence 188, App
110	26	92.9	6027	4	US-09-620-312D-517	Sequence 517, App	c 183	24	85.7	353	4	US-09-513-999C-1045	Sequence 1045, Ap
111	26	92.9	6444	4	US-08-956-171E-233	Sequence 233, App	c 184	24	85.7	357	4	US-09-248-796A-7762	Sequence 7762, Ap
112	26	92.9	6444	4	US-08-781-986A-233	Sequence 233, App	c 185	24	85.7	432	4	US-09-710-279-1713	Sequence 1713, Ap
113	26	92.9	9592	1	US-08-393-734-3	Sequence 3, Appli	c 186	24	85.7	439	4	US-09-513-999C-13645	Sequence 13645, A
114	26	92.9	9592	3	US-08-894-489-3	Sequence 3, Appli	c 187	24	85.7	450	4	US-09-513-999C-11985	Sequence 11985, A
115	26	92.9	12752	2	US-08-459-146-1	Sequence 1, Appli	c 188	24	85.7	458	4	US-09-513-999C-1912	Sequence 1912, Ap
116	26	92.9	12752	2	US-08-459-065-1	Sequence 1, Appli	c 189	24	85.7	463	4	US-09-270-767-2260	Sequence 2260, Ap
117	26	92.9	26000	4	US-09-843-376-10	Sequence 10, Appl	c 190	24	85.7	463	4	US-09-270-767-17542	Sequence 17542, A
118	26	92.9	44453	3	US-09-146-053-5	Sequence 5, Appli	191	24	85.7	483	4	US-09-621-976-3002	Sequence 3002, Ap
119	26	92.9	118999	4	US-09-791-105B-32	Sequence 32, Appl	192	24	85.7	523	4	US-09-621-976-14278	Sequence 14278, A
120	26	92.9	118999	4	US-09-791-105B-32	Sequence 32, Appl	193	24	85.7	585	4	US-09-252-991A-6323	Sequence 6323, Ap
121	26	92.9	129908	4	US-09-585-858-1	Sequence 1, Appli	194	24	85.7	603	4	US-09-134-000C-797	Sequence 797, App
122	26	92.9	152331	3	US-09-128-155-16	Sequence 16, Appl	195	24	85.7	614	3	US-09-385-982-18	Sequence 18, Appl
123	26	92.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli	c 196	24	85.7	666	4	US-10-101-464A-217	Sequence 217, App
124	26	92.9	1664976	4	US-09-692-570-1	Sequence 1, Appli	c 197	24	85.7	756	4	US-09-543-681A-1426	Sequence 1426, Ap
125	25	89.3	189	4	US-09-107-532A-901	Sequence 901, App	c 198	24	85.7	756	4	US-09-489-039A-1842	Sequence 1842, Ap
126	25	89.3	192	4	US-09-513-999C-17100	Sequence 17100, A	199	24	85.7	795	4	US-09-710-279-775	Sequence 775, App
127	25	89.3	249	4	US-09-252-991A-1965	Sequence 1965, Ap	200	24	85.7	807	3	US-09-134-001C-1892	Sequence 1892, Ap
128	25	89.3	342	4	US-09-513-999C-31025	Sequence 31025, A	201	24	85.7	855	4	US-09-252-991A-16249	Sequence 16249, A
129	25	89.3	423	4	US-09-270-767-1681	Sequence 1681, Ap	202	24	85.7	1005	4	US-09-489-039A-2157	Sequence 2157, Ap
130	25	89.3	423	4	US-09-270-767-16963	Sequence 16963, A	203	24	85.7	1056	4	US-09-107-532A-1039	Sequence 1039, Ap
131	25	89.3	525	4	US-09-107-532A-2325	Sequence 2325, Ap	204	24	85.7	1116	4	US-09-134-000C-2684	Sequence 2684, Ap
132	25	89.3	534	4	US-09-134-000C-4	Sequence 4, Appli	205	24	85.7	1130	4	US-09-386-653A-8	Sequence 8, Appli
133	25	89.3	574	4	US-09-621-976-9839	Sequence 9839, Ap	c 206	24	85.7	1211	4	US-09-270-767-13576	Sequence 13576, A
134	25	89.3	723	4	US-09-583-110-818	Sequence 818, App	207	24	85.7	1211	4	US-10-083-889-21	Sequence 21, Appl
135	25	89.3	936	4	US-09-393-858-19	Sequence 19, Appl	208	24	85.7	1530	4	US-10-083-889-1	Sequence 1, Appli
136	25	89.3	936	4	US-09-393-858-21	Sequence 21, Appl	209	24	85.7	1572	4	US-09-134-000C-2601	Sequence 2601, Ap
137	25	89.3	936	4	US-10-190-279-19	Sequence 19, Appl	210	24	85.7	1587	4	US-09-543-681A-3442	Sequence 3442, Ap
138	25	89.3	936	4	US-10-190-279-21	Sequence 21, Appl	211	24	85.7	1599	4	US-09-543-681A-3442	Sequence 3442, Ap
139	25	89.3	1038	4	US-09-252-991A-9759	Sequence 9759, Ap	c 212	24	85.7	1661	1	US-08-592-126-68	Sequence 68, Appl
140	25	89.3	1041	4	US-09-252-991A-9847	Sequence 9847, Ap	c 213	24	85.7	1661	4	US-09-168-595-68	Sequence 68, Appl
141	25	89.3	1446	4	US-09-107-532A-3258	Sequence 3258, Ap	c 214	24	85.7	1773	4	US-09-248-796A-12771	Sequence 12771, A
142	25	89.3	1452	4	US-09-489-039A-2735	Sequence 2735, Ap	215	24	85.7	1788	4	US-09-614-221A-394	Sequence 394, App
143	25	89.3	1569	4	US-09-489-039A-2870	Sequence 2870, Ap	c 216	24	85.7	1831	3	US-09-433-248A-5	Sequence 5, Appli
144	25	89.3	1783	4	US-09-221-017B-662	Sequence 662, App	217	24	85.7	1884	4	US-09-403-667A-1	Sequence 1, Appli
145	25	89.3	2541	4	US-09-252-991A-1828	Sequence 1828, Ap	218	24	85.7	1884	4	US-09-403-667A-3	Sequence 3, Appli
146	25	89.3	2610	4	US-09-489-039A-556	Sequence 556, App	219	24	85.7	1893	4	US-09-248-796A-5516	Sequence 5516, Ap
147	25	89.3	2670	4	US-09-252-991A-2131	Sequence 2131, Ap	220	24	85.7	1955	4	US-09-482-273-98	Sequence 98, Appl
148	25	89.3	2805	4	US-09-252-991A-10208	Sequence 10208, Ap	c 221	24	85.7	2034	4	US-09-252-991A-1531	Sequence 1531, Ap
149	25	89.3	3144	4	US-08-961-527-224	Sequence 224, App	222	24	85.7	2241	4	US-09-603-448-10	Sequence 10, Appl
150	25	89.3	3150	3	US-08-943-768-1	Sequence 1, Appli	c 223	24	85.7	2346	4	US-09-248-796A-5034	Sequence 5034, Ap
151	25	89.3	3150	4	US-09-865-960-1	Sequence 1, Appli	c 224	24	85.7	2586	4	US-10-101-464A-883	Sequence 883, App
152	25	89.3	3459	4	US-09-221-017B-257	Sequence 257, App	225	24	85.7	2635	3	US-09-186-276B-57	Sequence 57, Appl
153	25	89.3	3556	1	US-07-971-624E-1	Sequence 1, Appli	226	24	85.7	2635	4	US-08-842-445-57	Sequence 57, Appl
154	25	89.3	3572	1	US-07-971-624E-2	Sequence 2, Appli	227	24	85.7	2635	4	US-09-186-188B-57	Sequence 57, Appl
155	25	89.3	4041	1	US-08-147-812-4	Sequence 1, Appli	c 228	24	85.7	2823	4	US-09-704-611-4	Sequence 4, Appli
156	25	89.3	4110	3	US-09-123-708-1	Sequence 1, Appli	229	24	85.7	2971	4	US-09-482-273-37	Sequence 37, Appl
157	25	89.3	4110	3	US-09-123-624-1	Sequence 1, Appli	230	24	85.7	3040	4	US-09-710-279-3789	Sequence 3789, Ap
158	25	89.3	4165	1	US-08-147-812-6	Sequence 6, Appli	231	24	85.7	3288	4	US-09-489-039A-2096	Sequence 2096, Ap

C 378	23	82.1	924	4	US-09-393-858-18	Sequence 18, Appl	451	23	82.1	1347	4	US-09-583-110-94	Sequence 94, Appl
C 379	23	82.1	924	4	US-10-190-279-16	Sequence 16, Appl	C 452	23	82.1	1349	1	US-08-374-686-1	Sequence 1, Appl
C 380	23	82.1	924	4	US-10-190-279-18	Sequence 18, Appl	C 453	23	82.1	1380	4	US-09-252-991A-2247	Sequence 2247, Ap
C 381	23	82.1	925	3	US-09-230-380-7	Sequence 7, Appl	454	23	82.1	1395	4	US-09-248-796A-4084	Sequence 4084, Ap
C 382	23	82.1	939	1	US-08-592-411-14	Sequence 14, Appl	455	23	82.1	1396	4	US-09-710-279-4453	Sequence 4453, Ap
C 383	23	82.1	939	1	US-08-591-501-15	Sequence 15, Appl	456	23	82.1	1407	4	US-09-325-932A-107	Sequence 107, App
C 384	23	82.1	942	1	US-08-592-411-16	Sequence 16, Appl	457	23	82.1	1410	4	US-09-107-532A-2282	Sequence 2282, Ap
C 385	23	82.1	942	1	US-08-591-501-14	Sequence 14, Appl	458	23	82.1	1415	4	US-09-023-655-52	Sequence 52, Appl
C 386	23	82.1	954	4	US-09-543-681A-2089	Sequence 2089, Ap	C 459	23	82.1	1430	2	US-08-902-518A-1	Sequence 1, Appl
C 387	23	82.1	957	4	US-09-583-110-991	Sequence 991, App	C 460	23	82.1	1437	4	US-09-023-655-1029	Sequence 1029, Ap
C 388	23	82.1	957	4	US-09-270-767-33	Sequence 33, Appl	C 461	23	82.1	1452	1	US-08-276-943-1	Sequence 1, Appl
C 389	23	82.1	957	4	US-09-270-767-15315	Sequence 15315, A	C 462	23	82.1	1452	2	US-08-716-841-1	Sequence 1, Appl
C 390	23	82.1	960	4	US-09-270-767-26531	Sequence 26531, A	C 463	23	82.1	1452	4	US-09-856-640-1	Sequence 1, Appl
C 391	23	82.1	972	4	US-09-248-796A-4658	Sequence 4658, Ap	C 464	23	82.1	1476	4	US-09-602-787A-497	Sequence 497, App
C 392	23	82.1	990	4	US-09-134-000C-2833	Sequence 2833, Ap	C 465	23	82.1	1478	4	US-09-216-393B-7	Sequence 7, Appl
C 393	23	82.1	993	4	US-09-107-532A-3251	Sequence 3251, Ap	C 466	23	82.1	1497	4	US-09-489-039A-3622	Sequence 3622, Ap
C 394	23	82.1	1001	4	US-09-671-317-407	Sequence 407, App	C 467	23	82.1	1509	4	US-09-071-035-321	Sequence 321, App
C 395	23	82.1	1001	4	US-09-671-317-427	Sequence 427, App	C 468	23	82.1	1512	4	US-09-270-767-4316	Sequence 4316, Ap
C 396	23	82.1	1011	3	US-08-975-762-18	Sequence 18, Appl	C 469	23	82.1	1512	4	US-09-270-767-19598	Sequence 19598, A
C 397	23	82.1	1011	3	US-08-821-324-18	Sequence 18, Appl	C 470	23	82.1	1518	4	US-09-328-352-1697	Sequence 1697, Ap
C 398	23	82.1	1011	3	US-09-295-028-18	Sequence 18, Appl	C 471	23	82.1	1520	4	US-09-270-767-11022	Sequence 11022, A
C 399	23	82.1	1011	3	US-09-106-582-18	Sequence 18, Appl	C 472	23	82.1	1524	4	US-09-252-991A-14902	Sequence 14902, A
C 400	23	82.1	1011	4	US-09-328-352-1712	Sequence 1712, Ap	C 473	23	82.1	1572	4	US-09-252-991A-15025	Sequence 15025, A
C 401	23	82.1	1011	4	US-09-159-469-18	Sequence 18, Appl	C 474	23	82.1	1597	4	US-09-270-767-13899	Sequence 13899, A
C 402	23	82.1	1011	4	US-09-693-542-18	Sequence 18, Appl	C 475	23	82.1	1614	4	US-09-252-991A-13021	Sequence 13021, A
C 403	23	82.1	1017	4	US-09-489-039A-12046	Sequence 3001, Ap	C 476	23	82.1	1633	1	US-08-441-629-3	Sequence 3, Appl
C 404	23	82.1	1018	4	US-09-270-767-12041	Sequence 2119, Ap	C 477	23	82.1	1633	3	US-08-776-207-3	Sequence 3, Appl
C 405	23	82.1	1026	4	US-09-134-000C-2119	Sequence 2150, A	C 478	23	82.1	1633	5	US-09-507-773-3	Sequence 3, Appl
C 406	23	82.1	1026	4	US-09-270-767-12503	Sequence 1135, Ap	C 479	23	82.1	1633	3	PCT-US95-09172-3	Sequence 3, Appl
C 407	23	82.1	1035	4	US-09-602-787A-551	Sequence 551, App	C 480	23	82.1	1651	3	US-09-065-999-5	Sequence 5, Appl
C 408	23	82.1	1038	4	US-09-489-039A-1135	Sequence 1135, Ap	C 481	23	82.1	1651	3	US-09-065-999-6	Sequence 6, Appl
C 409	23	82.1	1047	4	US-09-270-767-29661	Sequence 29661, A	C 482	23	82.1	1660	4	US-09-221-017B-461	Sequence 461, App
C 410	23	82.1	1056	1	US-08-785-052-1	Sequence 1, Appl	C 483	23	82.1	1692	3	US-09-134-001C-79	Sequence 79, Appl
C 411	23	82.1	1056	2	US-08-913-581-1	Sequence 1, Appl	C 484	23	82.1	1738	4	US-09-227-421-4	Sequence 4, Appl
C 412	23	82.1	1062	3	US-08-838-151A-48	Sequence 48, Appl	C 485	23	82.1	1738	4	US-09-479-855-4	Sequence 4, Appl
C 413	23	82.1	1062	3	US-08-838-151A-51	Sequence 51, Appl	C 486	23	82.1	1782	4	US-09-252-991A-3870	Sequence 3870, Ap
C 414	23	82.1	1062	3	US-08-838-151A-54	Sequence 54, Appl	C 487	23	82.1	1809	3	US-09-134-001C-1579	Sequence 1579, Ap
C 415	23	82.1	1062	6	5395759-1	Patent No. 5395759	C 488	23	82.1	1836	4	US-09-252-991A-4587	Sequence 4587, Ap
C 416	23	82.1	1071	4	US-09-328-352-2132	Sequence 2132, Ap	C 489	23	82.1	1846	4	US-09-516-277-5	Sequence 5, Appl
C 417	23	82.1	1077	4	US-09-252-991A-3983	Sequence 3983, Ap	C 490	23	82.1	1887	4	US-09-107-532A-2783	Sequence 2783, Ap
C 418	23	82.1	1086	3	US-09-134-001C-1025	Sequence 1025, Ap	C 491	23	82.1	1894	3	US-09-065-999-8	Sequence 8, Appl
C 419	23	82.1	1135	4	US-09-221-017B-224	Sequence 224, App	C 492	23	82.1	1907	4	US-09-270-767-15061	Sequence 15061, A
C 420	23	82.1	1140	4	US-09-252-991A-3205	Sequence 3205, A	C 493	23	82.1	1947	4	US-09-252-991A-14728	Sequence 14728, A
C 421	23	82.1	1163	4	US-09-270-767-13507	Sequence 13507, A	C 494	23	82.1	1977	4	US-09-252-991A-3924	Sequence 3924, Ap
C 422	23	82.1	1173	2	US-08-184-009-214	Sequence 214, App	C 495	23	82.1	1991	4	US-09-270-767-11487	Sequence 11487, A
C 423	23	82.1	1173	2	US-08-458-356-214	Sequence 214, App	C 496	23	82.1	2001	4	US-09-248-796A-4560	Sequence 4560, Ap
C 424	23	82.1	1173	3	US-08-796-101-45	Sequence 45, Appl	C 497	23	82.1	2019	4	US-09-409-096-3	Sequence 3, Appl
C 425	23	82.1	1173	3	US-08-460-736-214	Sequence 214, App	C 498	23	82.1	2028	4	US-09-248-796A-6596	Sequence 6596, Ap
C 426	23	82.1	1173	3	US-09-485-885-13	Sequence 13, Appl	C 499	23	82.1	2029	3	US-09-232-160-13	Sequence 13, Appl
C 427	23	82.1	1173	4	US-09-535-370-214	Sequence 214, App	C 500	23	82.1	2040	4	US-09-252-991A-16099	Sequence 16099, A
C 428	23	82.1	1173	4	US-09-663-667-214	Sequence 214, App	C 501	23	82.1	2061	4	US-09-252-991A-2761	Sequence 2761, Ap
C 429	23	82.1	1178	4	US-09-952-689-4	Sequence 4, Appl	C 502	23	82.1	2072	3	US-09-065-999-7	Sequence 7, Appl
C 430	23	82.1	1179	4	US-09-489-039A-4085	Sequence 4085, Ap	C 503	23	82.1	2091	4	US-09-252-991A-4016	Sequence 4016, Ap
C 431	23	82.1	1183	3	US-08-838-151A-43	Sequence 43, Appl	C 504	23	82.1	2103	4	US-09-270-767-13648	Sequence 13648, A
C 432	23	82.1	1183	3	US-08-838-151A-45	Sequence 45, Appl	C 505	23	82.1	2107	3	US-09-180-852-1	Sequence 1, Appl
C 433	23	82.1	1188	4	US-09-270-767-12519	Sequence 12519, A	C 506	23	82.1	2148	4	US-09-237-421-1	Sequence 1, Appl
C 434	23	82.1	1191	3	US-09-283-305-13	Sequence 13, Appl	C 507	23	82.1	2148	4	US-09-479-855-1	Sequence 1, Appl
C 435	23	82.1	1191	4	US-09-883-720-13	Sequence 13, Appl	C 508	23	82.1	2154	4	US-09-134-000C-2428	Sequence 2428, Ap
C 436	23	82.1	1224	4	US-09-252-991A-2437	Sequence 2437, Ap	C 509	23	82.1	2202	4	US-09-248-796A-2462	Sequence 2462, Ap
C 437	23	82.1	1226	1	PCT-US91-00909-3	Sequence 3, Appl	C 510	23	82.1	2203	4	US-09-801-861-1	Sequence 1, Appl
C 438	23	82.1	1242	5	PCT-US91-00909-3	Sequence 3, Appl	C 511	23	82.1	2203	4	US-10-224-562-1	Sequence 1, Appl
C 439	23	82.1	1245	3	US-09-283-305-15	Sequence 15, Appl	C 512	23	82.1	2211	4	US-09-107-532A-2353	Sequence 2353, Ap
C 440	23	82.1	1245	3	US-09-883-720-15	Sequence 15, Appl	C 513	23	82.1	2215	6	5457037-2	Patent No. 5457037
C 441	23	82.1	1250	4	US-09-710-279-949	Sequence 949, App	C 514	23	82.1	2256	4	US-09-252-991A-16502	Sequence 16502, A
C 442	23	82.1	1254	4	US-09-312-283C-67	Sequence 67, Appl	C 515	23	82.1	2262	4	US-09-248-796A-4938	Sequence 4938, Ap
C 443	23	82.1	1283	3	US-09-283-305-11	Sequence 11, Appl	C 516	23	82.1	2305	4	US-09-016-434-1282	Sequence 1282, Ap
C 444	23	82.1	1283	4	US-09-883-720-11	Sequence 11, Appl	C 517	23	82.1	2305	4	US-09-023-655-1249	Sequence 1249, Ap
C 445	23	82.1	1291	4	US-09-976-594-598	Sequence 588, App	C 518	23	82.1	2309	4	US-09-855-323-9	Sequence 9, Appl
C 446	23	82.1	1293	4	US-09-248-796A-5389	Sequence 5389, Ap	C 519	23	82.1	2322	3	US-08-975-762-44	Sequence 44, Appl
C 447	23	82.1	1308	4	US-09-801-861-4	Sequence 4, Appl	C 520	23	82.1	2322	3	US-09-295-028-44	Sequence 44, Appl
C 448	23	82.1	1308	4	US-10-224-562-4	Sequence 4, Appl	C 521	23	82.1	2322	3	US-09-106-582-44	Sequence 44, Appl
C 449	23	82.1	1317	4	US-09-248-796A-9816	Sequence 9816, Ap	C 522	23	82.1	2322	4	US-09-159-469-44	Sequence 44, Appl
C 450	23	82.1	1332	4	US-09-270-767-11937	Sequence 11937, A	C 523	23	82.1	2322	4	US-09-693-542-44	Sequence 44, Appl

524	23	82.1	2343	4	US-09-484-970B-119	Sequence 119, App	c 597	23	82.1	3969	1	US-08-818-823-11	Sequence 11, Appl
525	23	82.1	2372	4	US-09-907-794A-200	Sequence 200, App	c 598	23	82.1	4010	5	PCT-US94-07297-40	Sequence 40, Appl
526	23	82.1	2372	4	US-09-905-125A-200	Sequence 200, App	c 599	23	82.1	4074	4	US-09-252-991A-4737	Sequence 4737, Ap
527	23	82.1	2372	4	US-09-902-775A-200	Sequence 200, App	600	23	82.1	4092	4	US-09-252-991A-4771	Sequence 4771, Ap
528	23	82.1	2372	4	US-09-906-700-200	Sequence 200, App	601	23	82.1	4104	3	US-08-881-706-1	Sequence 1, Appli
529	23	82.1	2372	4	US-09-903-603A-200	Sequence 200, App	602	23	82.1	4104	4	US-09-823-394-1	Sequence 1, Appli
c 530	23	82.1	2373	4	US-09-543-681A-1675	Sequence 1675, Ap	c 603	23	82.1	4131	1	US-07-872-644-38	Sequence 38, Appl
c 531	23	82.1	2406	1	US-08-396-479B-7	Sequence 7, Appli	c 604	23	82.1	4131	1	US-08-297-494-38	Sequence 38, Appl
c 532	23	82.1	2406	1	US-08-818-823-7	Sequence 7, Appli	c 605	23	82.1	4131	1	US-08-297-510-38	Sequence 38, Appl
c 533	23	82.1	2472	4	US-09-248-796A-5236	Sequence 5236, Ap	c 606	23	82.1	4131	1	US-08-479-532-38	Sequence 38, Appl
c 534	23	82.1	2504	1	US-08-121-713D-63	Sequence 63, Appl	c 607	23	82.1	4131	1	US-08-455-526-38	Sequence 38, Appl
c 535	23	82.1	2504	1	US-08-835-268-63	Sequence 63, Appl	c 608	23	82.1	4131	1	US-08-455-526-38	Sequence 38, Appl
c 536	23	82.1	2504	2	US-09-060-692-63	Sequence 63, Appl	c 609	23	82.1	4131	3	US-09-133-451-38	Sequence 38, Appl
c 537	23	82.1	2504	3	US-08-833-391-63	Sequence 63, Appl	c 610	23	82.1	4131	4	US-09-883-825-38	Sequence 38, Appl
c 538	23	82.1	2504	3	US-09-060-610-63	Sequence 63, Appl	c 611	23	82.1	4131	5	PCT-US92-03222-38	Sequence 38, Appl
c 539	23	82.1	2504	5	PCT-US94-10151A-63	Sequence 63, Appl	c 612	23	82.1	4171	3	US-09-754-250-1	Sequence 1, Appli
540	23	82.1	2520	3	US-08-968-563-10	Sequence 10, Appl	613	23	82.1	4240	4	US-09-708-392-6	Sequence 6, Appli
541	23	82.1	2520	3	US-08-969-683A-10	Sequence 10, Appl	614	23	82.1	4297	4	US-09-710-279-4092	Sequence 4092, Ap
542	23	82.1	2520	3	US-09-297-928-6	Sequence 6, Appli	c 615	23	82.1	4408	4	US-09-221-017B-830	Sequence 830, App
543	23	82.1	2550	2	US-08-884-072-2	Sequence 2, Appli	616	23	82.1	4833	3	US-09-066-047-1	Sequence 1, Appli
544	23	82.1	2550	3	US-09-212-168-2	Sequence 2, Appli	617	23	82.1	5021	4	US-09-285-385C-1	Sequence 1, Appli
545	23	82.1	2589	4	US-09-614-221A-389	Sequence 389, App	618	23	82.1	5077	3	US-09-480-921B-24	Sequence 24, Appl
c 546	23	82.1	2616	4	US-09-252-991A-2860	Sequence 2860, Ap	619	23	82.1	5510	4	US-08-956-171E-165	Sequence 165, App
547	23	82.1	2617	4	US-09-799-978-25	Sequence 25, Appl	620	23	82.1	5510	4	US-08-781-986A-165	Sequence 165, App
548	23	82.1	2625	6	5457037-4	Patent No. 5457037	621	23	82.1	5519	3	US-08-737-607-1	Sequence 1, Appli
549	23	82.1	2643	2	US-08-750-134A-10	Sequence 10, Appl	622	23	82.1	5521	3	US-08-975-762-48	Sequence 48, Appl
550	23	82.1	2643	3	US-09-363-745-10	Sequence 10, Appl	623	23	82.1	5521	3	US-09-295-028-48	Sequence 48, Appl
551	23	82.1	2643	3	US-09-023-655-897	Sequence 897, App	624	23	82.1	5521	3	US-09-106-582-48	Sequence 48, Appl
c 552	23	82.1	2647	1	US-08-396-479B-9	Sequence 9, Appli	625	23	82.1	5521	4	US-09-159-469-48	Sequence 48, Appl
c 553	23	82.1	2647	1	US-08-818-823-9	Sequence 9, Appli	626	23	82.1	5521	4	US-08-693-542-48	Sequence 48, Appl
c 554	23	82.1	2675	4	US-09-710-279-4387	Sequence 4387, Ap	c 627	23	82.1	5643	1	US-08-144-602B-4	Sequence 4, Appli
555	23	82.1	2679	4	US-09-328-352-4038	Sequence 4038, Ap	628	23	82.1	5666	4	US-10-204-708-30	Sequence 30, Appl
556	23	82.1	2691	3	US-09-453-702B-227	Sequence 227, App	c 629	23	82.1	5829	3	US-08-809-254A-4	Sequence 4, Appli
c 557	23	82.1	2694	4	US-09-252-991A-12457	Sequence 12457, A	630	23	82.1	5857	3	US-09-293-170-4	Sequence 4, Appli
c 558	23	82.1	2735	4	US-10-101-464A-865	Sequence 865, App	c 631	23	82.1	5925	4	US-09-315-926A-78	Sequence 78, Appl
c 559	23	82.1	2799	4	US-09-181-339-2	Sequence 2, Appli	c 632	23	82.1	5925	4	US-10-204-708-25	Sequence 25, Appl
c 560	23	82.1	2823	4	US-10-140-002-183	Sequence 183, App	c 633	23	82.1	7304	3	US-09-453-702B-174	Sequence 174, App
561	23	82.1	2847	1	US-07-747-901A-2	Sequence 2, Appli	634	23	82.1	7507	2	US-08-975-763-1	Sequence 1, Appli
562	23	82.1	2847	1	US-07-935-312-2	Sequence 2, Appli	635	23	82.1	7650	4	US-09-221-017B-911	Sequence 911, App
563	23	82.1	2847	4	US-09-328-352-1997	Sequence 1997, Ap	636	23	82.1	7714	4	US-09-479-132-26	Sequence 26, Appl
564	23	82.1	2860	4	US-09-991-458A-1	Sequence 1, Appli	637	23	82.1	7714	4	US-09-484-997-26	Sequence 26, Appl
c 565	23	82.1	2898	4	US-09-489-039A-1679	Sequence 1679, Ap	638	23	82.1	7714	4	US-09-481-355-26	Sequence 26, Appl
c 566	23	82.1	2929	1	US-09-710-279-3857	Sequence 3857, Ap	639	23	82.1	7714	4	US-09-481-282-26	Sequence 26, Appl
c 567	23	82.1	2935	1	US-08-246-583-1	Sequence 1, Appli	640	23	82.1	7714	4	US-09-485-659A-26	Sequence 26, Appl
c 568	23	82.1	2940	4	US-09-636-791A-3	Sequence 3, Appli	641	23	82.1	7714	4	US-09-484-996-26	Sequence 26, Appl
c 569	23	82.1	3001	4	US-09-539-333D-132	Sequence 132, App	642	23	82.1	7714	4	US-09-479-123-26	Sequence 26, Appl
570	23	82.1	3065	3	US-09-199-637A-129	Sequence 129, App	643	23	82.1	7714	4	US-09-484-317A-26	Sequence 26, Appl
571	23	82.1	3116	1	US-08-149-103-2	Sequence 2, Appli	644	23	82.1	8537	4	US-10-204-708-42	Sequence 42, Appl
572	23	82.1	3116	1	US-08-451-883-2	Sequence 2, Appli	645	23	82.1	9828	4	US-08-961-527-41	Sequence 41, Appl
c 573	23	82.1	3137	4	US-09-710-279-3842	Sequence 3842, Ap	646	23	82.1	10254	3	US-09-058-746-1	Sequence 1, Appli
574	23	82.1	3168	1	US-07-723-002C-7	Sequence 7, Appli	647	23	82.1	10317	3	US-09-438-142-1	Sequence 1, Appli
575	23	82.1	3314	4	US-09-536-882A-5	Sequence 5, Appli	648	23	82.1	10317	3	US-09-438-142-1	Sequence 1, Appli
c 576	23	82.1	3326	4	US-09-710-279-3952	Sequence 3952, Ap	649	23	82.1	10660	2	US-08-267-803B-8	Sequence 8, Appli
c 577	23	82.1	3336	6	5457037-1	Patent No. 5457037	650	23	82.1	10660	3	US-09-041-886B-16	Sequence 16, Appl
c 578	23	82.1	3390	4	US-09-252-991A-11981	Sequence 11981, A	c 651	23	82.1	11236	1	US-07-853-913-1	Sequence 1, Appli
c 579	23	82.1	3396	3	US-08-974-549A-639	Sequence 639, App	652	23	82.1	11474	4	US-09-495-406-1	Sequence 1, Appli
c 580	23	82.1	3396	4	US-09-721-456-639	Sequence 639, App	653	23	82.1	11474	4	US-09-816-028A-1	Sequence 1, Appli
581	23	82.1	3449	4	US-09-221-017B-257	Sequence 257, App	654	23	82.1	11474	4	US-10-303-162-1	Sequence 1, Appli
c 582	23	82.1	3451	3	US-08-974-549A-721	Sequence 721, App	655	23	82.1	12597	4	US-09-705-299-12	Sequence 12, Appl
c 583	23	82.1	3451	4	US-09-811-286-1	Sequence 1, Appli	656	23	82.1	14872	4	US-08-961-527-72	Sequence 72, Appl
c 584	23	82.1	3451	4	US-09-721-456-721	Sequence 721, App	657	23	82.1	14872	4	US-10-204-708-39	Sequence 39, Appl
c 585	23	82.1	3461	4	US-10-140-002-139	Sequence 139, App	c 658	23	82.1	19513	4	US-08-956-171E-38	Sequence 38, Appl
c 586	23	82.1	3569	4	US-09-710-279-4443	Sequence 4443, Ap	c 659	23	82.1	23439	4	US-08-781-986A-38	Sequence 38, Appl
587	23	82.1	3671	4	US-09-221-017B-1111	Sequence 1111, Ap	c 660	23	82.1	23451	3	US-09-453-702B-173	Sequence 173, App
c 588	23	82.1	3680	4	US-09-023-655-1477	Sequence 1477, Ap	661	23	82.1	26700	1	US-08-472-217-1	Sequence 1, Appli
c 589	23	82.1	3729	4	US-09-710-279-3675	Sequence 3675, Ap	662	23	82.1	26700	2	US-08-488-139-5	Sequence 5, Appli
c 590	23	82.1	3797	3	US-08-936-165A-107	Sequence 107, App	663	23	82.1	26700	3	US-08-760-534A-1	Sequence 1, Appli
c 591	23	82.1	3864	4	US-09-710-279-3727	Sequence 3727, Ap	664	23	82.1	26700	4	US-09-336-757-1	Sequence 1, Appli
592	23	82.1	3879	4	US-09-638-715-35	Sequence 35, Appl	665	23	82.1	40090	4	US-09-820-004-3	Sequence 3, Appli
593	23	82.1	3879	4	US-10-060-509-35	Sequence 35, Appl	666	23	82.1	43676	3	US-09-356-952-12	Sequence 12, Appl
594	23	82.1	3879	4	US-10-060-506-35	Sequence 35, Appl	667	23	82.1	47981	4	US-09-679-279-1	Sequence 1, Appli
595	23	82.1	3915	4	US-09-023-655-1104	Sequence 1104, Ap	668	23	82.1	50341	2	US-08-247-901C-1	Sequence 1, Appli
c 596	23	82.1	3969	1	US-08-396-479B-11	Sequence 11, Appl	669	23	82.1	50341	2	US-09-075-904-1	Sequence 1, Appli

670	23	82.1	52297	3	US-09-426-436-1	Sequence 1, Appli	743	22	78.6	283	4	US-09-313-294A-3793	Sequence 3793, Ap
671	23	82.1	52297	3	US-08-705-557-1	Sequence 1, Appli	744	22	78.6	285	4	US-09-724-623-46	Sequence 46, Appli
672	23	82.1	70000	4	US-09-851-896-3	Sequence 3, Appli	745	22	78.6	285	4	US-08-844-023-4	Sequence 4, Appli
673	23	82.1	84495	3	US-09-797-906-3	Sequence 3, Appli	c 746	22	78.6	285	4	US-09-513-999C-26667	Sequence 26667, A
674	23	82.1	92407	4	US-09-596-002-36	Sequence 36, Appli	747	22	78.6	286	2	US-08-447-173A-59	Sequence 59, Appli
675	23	82.1	99500	3	US-09-799-096-10	Sequence 10, Appli	748	22	78.6	286	2	US-08-447-173A-60	Sequence 60, Appli
676	23	82.1	111282	3	US-09-754-250-3	Sequence 3, Appli	749	22	78.6	289	4	US-09-313-294A-5707	Sequence 5707, Ap
677	23	82.1	137000	4	US-10-172-911-11	Sequence 11, Appli	c 750	22	78.6	291	4	US-09-270-767-28709	Sequence 28709, A
678	23	82.1	148567	4	US-09-801-876B-3	Sequence 3, Appli	c 751	22	78.6	292	4	US-09-720-655B-3	Sequence 3, Appli
679	23	82.1	148567	4	US-10-254-869-3	Sequence 3, Appli	752	22	78.6	295	4	US-09-313-294A-2901	Sequence 2901, Ap
c 680	23	82.1	202001	4	US-09-734-674-3	Sequence 3, Appli	c 753	22	78.6	297	4	US-09-252-991A-9993	Sequence 9993, Ap
681	23	82.1	246240	2	US-08-724-394A-20	Sequence 20, Appli	c 754	22	78.6	297	4	US-09-134-000C-236	Sequence 236, App
682	23	82.1	246240	2	US-08-724-394A-21	Sequence 21, Appli	755	22	78.6	299	4	US-09-513-999C-12691	Sequence 12691, A
683	23	82.1	246240	2	US-08-724-394A-22	Sequence 22, Appli	756	22	78.6	303	4	US-09-513-999C-11097	Sequence 11097, A
684	23	82.1	536165	4	US-09-214-808-1	Sequence 1, Appli	757	22	78.6	305	4	US-09-313-294A-4746	Sequence 4746, Ap
685	23	82.1	580073	4	US-08-545-528D-1	Sequence 1, Appli	758	22	78.6	308	4	US-08-844-023-3	Sequence 3, Appli
c 686	23	82.1	580073	4	US-08-545-528D-1	Sequence 1, Appli	759	22	78.6	314	1	US-09-609-324A-7	Sequence 7, Appli
c 687	23	82.1	1230025	4	US-09-198-452A-1	Sequence 1, Appli	760	22	78.6	314	2	US-08-920-440B-7	Sequence 7, Appli
c 688	23	82.1	1830121	4	US-09-557-884-1	Sequence 1, Appli	761	22	78.6	314	3	US-09-173-492-7	Sequence 7, Appli
c 689	23	82.1	1830121	4	US-09-643-990A-1	Sequence 1, Appli	762	22	78.6	314	3	US-09-173-133-7	Sequence 7, Appli
c 690	23	82.1	1830121	4	US-10-329-960-1	Sequence 1, Appli	763	22	78.6	314	3	US-09-165-533-7	Sequence 7, Appli
c 691	23	82.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli	764	22	78.6	314	4	US-09-580-236A-7	Sequence 7, Appli
c 692	23	82.1	441529	3	US-09-103-840A-1	Sequence 1, Appli	765	22	78.6	318	4	US-09-221-017B-458	Sequence 458, App
c 693	22	78.6	16	2	US-08-447-173A-64	Sequence 64, Appli	766	22	78.6	318	4	US-09-513-999C-1396	Sequence 1396, Ap
c 694	22	78.6	20	4	US-09-723-368-8	Sequence 8, Appli	767	22	78.6	322	3	US-09-323-873A-9	Sequence 9, Appli
c 695	22	78.6	21	4	US-09-380-836-55	Sequence 55, Appli	768	22	78.6	326	4	US-08-956-171E-4450	Sequence 4450, Ap
c 696	22	78.6	21	4	US-09-380-836-56	Sequence 56, Appli	769	22	78.6	326	4	US-08-781-986A-4450	Sequence 4450, Ap
c 697	22	78.6	32	3	US-09-230-288-13	Sequence 13, Appli	c 770	22	78.6	330	1	US-08-318-970B-36	Sequence 36, Appli
c 698	22	78.6	48	2	US-08-865-675-7	Sequence 7, Appli	c 771	22	78.6	330	1	US-08-318-970B-37	Sequence 37, Appli
c 699	22	78.6	48	2	US-09-237-510-7	Sequence 7, Appli	c 772	22	78.6	330	1	US-08-442-542-7	Sequence 7, Appli
c 700	22	78.6	66	4	US-09-513-999C-34862	Sequence 34862, A	c 773	22	78.6	330	3	US-08-483-749A-19	Sequence 19, Appli
c 701	22	78.6	75	4	US-08-956-171E-3338	Sequence 3338, Ap	c 774	22	78.6	330	3	US-08-765-469-7	Sequence 7, Appli
c 702	22	78.6	75	4	US-08-781-986A-3338	Sequence 3338, Ap	775	22	78.6	332	4	US-10-012-282-3	Sequence 3, Appli
c 703	22	78.6	81	1	US-08-458-023B-14	Sequence 14, Appli	c 776	22	78.6	336	1	US-08-318-970B-35	Sequence 35, Appli
c 704	22	78.6	95	2	US-08-353-476-61	Sequence 61, Appli	c 777	22	78.6	339	4	US-09-252-991A-12495	Sequence 12495, A
705	22	78.6	100	3	US-09-298-886-14	Sequence 14, Appli	c 778	22	78.6	342	4	US-09-134-000C-1407	Sequence 1407, Ap
706	22	78.6	100	3	US-09-298-886-17	Sequence 17, Appli	c 779	22	78.6	348	4	US-09-252-991A-12495	Sequence 12495, A
c 707	22	78.6	100	4	US-09-621-976-13894	Sequence 13894, A	c 780	22	78.6	348	4	US-09-134-000C-2798	Sequence 2798, Ap
c 708	22	78.6	100	4	US-09-993-672-14	Sequence 14, Appli	781	22	78.6	351	4	US-09-489-039A-4447	Sequence 4447, Ap
c 709	22	78.6	100	4	US-09-993-672-17	Sequence 17, Appli	c 782	22	78.6	354	1	US-08-318-970B-44	Sequence 44, Appli
c 710	22	78.6	111	3	US-08-767-128-39	Sequence 39, Appli	c 783	22	78.6	354	4	US-09-134-000C-2171	Sequence 2171, Ap
711	22	78.6	112	4	US-08-956-171E-1527	Sequence 1527, Ap	c 784	22	78.6	366	3	US-09-030-607-215	Sequence 215, App
712	22	78.6	112	4	US-08-781-986A-1527	Sequence 1527, Ap	c 785	22	78.6	366	3	US-09-439-313-215	Sequence 215, App
713	22	78.6	166	4	US-08-956-171E-4703	Sequence 4703, Ap	c 786	22	78.6	366	3	US-09-352-616A-215	Sequence 215, App
714	22	78.6	166	4	US-08-781-986A-4703	Sequence 4703, Ap	c 787	22	78.6	366	4	US-09-232-149A-215	Sequence 215, App
715	22	78.6	183	4	US-09-107-522A-870	Sequence 870, App	c 788	22	78.6	366	4	US-09-489-039A-3977	Sequence 3977, Ap
716	22	78.6	189	4	US-09-328-352-2107	Sequence 2107, Ap	c 789	22	78.6	366	4	US-09-159-812-215	Sequence 215, App
c 717	22	78.6	193	3	US-09-063-733A-53	Sequence 53, Appli	c 790	22	78.6	366	4	US-09-636-215-215	Sequence 215, App
c 718	22	78.6	202	4	US-09-016-434-285	Sequence 285, App	c 791	22	78.6	366	4	US-09-685-166A-215	Sequence 215, App
719	22	78.6	205	3	US-09-218-363-27	Sequence 27, Appli	c 792	22	78.6	366	4	US-09-115-453-215	Sequence 215, App
c 720	22	78.6	206	4	US-09-513-999C-35499	Sequence 35499, A	c 793	22	78.6	366	4	US-09-688-489-215	Sequence 215, App
c 721	22	78.6	210	2	US-08-471-227-1	Sequence 1, Appli	c 794	22	78.6	366	4	US-09-679-426-215	Sequence 215, App
c 722	22	78.6	222	4	US-09-489-039A-1949	Sequence 1949, Ap	c 795	22	78.6	367	4	US-09-270-767-394	Sequence 394, App
c 723	22	78.6	225	4	US-09-134-000C-2517	Sequence 2517, Ap	c 796	22	78.6	367	4	US-09-270-767-15676	Sequence 15676, A
724	22	78.6	228	4	US-09-134-000C-1406	Sequence 1406, Ap	c 797	22	78.6	370	3	US-09-450-656-2	Sequence 2, Appli
725	22	78.6	243	4	US-09-313-294A-5621	Sequence 5621, Ap	798	22	78.6	372	2	US-08-454-557C-35	Sequence 35, Appli
c 726	22	78.6	245	4	US-09-513-999C-23892	Sequence 23892, A	799	22	78.6	372	2	US-08-340-426D-35	Sequence 35, Appli
c 727	22	78.6	246	4	US-09-702-705-1650	Sequence 1650, Ap	800	22	78.6	372	2	US-08-450-673C-35	Sequence 35, Appli
c 728	22	78.6	246	4	US-09-736-457-1650	Sequence 1650, Ap	801	22	78.6	372	5	PCT-US95-17111A-35	Sequence 35, Appli
c 729	22	78.6	246	4	US-09-614-124B-1650	Sequence 1650, Ap	802	22	78.6	385	4	US-09-513-999C-34497	Sequence 34497, A
c 730	22	78.6	246	4	US-09-671-325-1650	Sequence 1650, Ap	803	22	78.6	393	4	US-09-134-000C-1213	Sequence 1213, Ap
731	22	78.6	246	4	US-09-583-110-359	Sequence 359, App	804	22	78.6	396	4	US-09-389-681-212	Sequence 1214, Ap
c 732	22	78.6	246	4	US-09-658-824-1650	Sequence 1650, Ap	805	22	78.6	401	4	US-09-620-405B-212	Sequence 212, App
733	22	78.6	248	4	US-09-513-999C-2740	Sequence 2740, Ap	806	22	78.6	401	4	US-09-339-338-212	Sequence 212, App
734	22	78.6	252	4	US-09-248-796A-13683	Sequence 13683, A	807	22	78.6	401	4	US-09-433-826B-212	Sequence 212, App
735	22	78.6	253	3	US-09-218-363-21	Sequence 21, Appli	808	22	78.6	401	4	US-09-604-287A-212	Sequence 212, App
c 736	22	78.6	253	4	US-09-513-999C-2937	Sequence 2937, Ap	809	22	78.6	401	4	US-08-834-759-212	Sequence 212, App
737	22	78.6	257	4	US-09-016-434-480	Sequence 480, App	810	22	78.6	401	4	US-09-590-751A-212	Sequence 212, App
c 738	22	78.6	264	4	US-09-313-294A-6240	Sequence 6240, Ap	811	22	78.6	401	4	US-09-621-976-14332	Sequence 14332, A
739	22	78.6	264	4	US-09-513-999C-34216	Sequence 34216, A	812	22	78.6	402	4	US-09-621-976-14332	Sequence 14332, A
740	22	78.6	273	4	US-09-696-569-1	Sequence 1, Appli	c 813	22	78.6	405	2	US-08-621-751A-11	Sequence 11, Appli
741	22	78.6	277	4	US-09-513-999C-28003	Sequence 28003, A	814	22	78.6	407	4	US-09-513-999C-27388	Sequence 27388, A
742	22	78.6	281	4	US-09-313-294A-6406	Sequence 6406, Ap	815	22	78.6	414	4	US-09-489-039A-1457	Sequence 1457, Ap

C 816	22	78.6	417	1	US-08-318-970B-46	Sequence 46, Appl	C 889	22	78.6	597	4	US-09-621-976-3245	Sequence 3245, Ap
C 817	22	78.6	418	4	US-09-621-976-13244	Sequence 13244, A	890	22	78.6	598	4	US-09-513-999C-3977	Sequence 3977, Ap
C 818	22	78.6	419	4	US-09-976-594-512	Sequence 512, App	C 891	22	78.6	603	4	US-09-252-991A-6776	Sequence 6776, Ap
C 819	22	78.6	420	4	US-09-484-970B-155	Sequence 155, App	892	22	78.6	603	4	US-09-248-796A-5574	Sequence 5574, Ap
C 820	22	78.6	421	4	US-09-134-000C-749	Sequence 749, App	893	22	78.6	609	4	US-09-252-991A-14745	Sequence 14745, A
C 821	22	78.6	422	4	US-09-252-991A-3928	Sequence 3928, Ap	894	22	78.6	612	4	US-09-252-991A-1623	Sequence 1623, Ap
C 822	22	78.6	423	4	US-09-621-976-17465	Sequence 17465, Ap	895	22	78.6	612	4	US-09-134-000C-2239	Sequence 2239, Ap
C 823	22	78.6	424	4	US-09-270-767-5167	Sequence 5167, Ap	896	22	78.6	612	4	US-09-270-767-8706	Sequence 8706, Ap
C 824	22	78.6	425	4	US-09-270-767-5614	Sequence 5614, Ap	C 897	22	78.6	612	4	US-09-270-767-23988	Sequence 23988, A
C 825	22	78.6	426	4	US-09-270-767-20449	Sequence 20449, A	898	22	78.6	615	4	US-09-270-767-1189	Sequence 1189, Ap
C 826	22	78.6	427	4	US-09-107-532A-1058	Sequence 1058, Ap	899	22	78.6	615	4	US-09-270-767-16471	Sequence 16471, A
C 827	22	78.6	428	4	US-09-107-532A-1058	Sequence 1058, Ap	900	22	78.6	616	4	US-09-221-017B-511	Sequence 511, App
C 828	22	78.6	429	4	US-09-621-976-10373	Sequence 10373, A	901	22	78.6	618	4	US-09-252-991A-4150	Sequence 4150, Ap
C 829	22	78.6	430	4	US-09-513-999C-25864	Sequence 25864, A	C 902	22	78.6	618	4	US-09-252-991A-15008	Sequence 15008, A
C 830	22	78.6	431	4	US-09-252-991A-16414	Sequence 16414, A	C 903	22	78.6	630	3	US-08-998-416-301	Sequence 301, App
C 831	22	78.6	432	4	US-08-318-970B-45	Sequence 45, Appl	904	22	78.6	641	4	US-09-615-192A-185	Sequence 185, App
C 832	22	78.6	433	2	US-08-653-402B-11	Sequence 11, Appl	905	22	78.6	642	1	US-09-609-324A-9	Sequence 9, Appl
C 833	22	78.6	434	4	US-09-621-976-8402	Sequence 8402, Ap	906	22	78.6	642	3	US-08-920-440B-9	Sequence 9, Appl
C 834	22	78.6	435	4	US-09-621-976-17333	Sequence 17333, A	907	22	78.6	642	3	US-09-173-432-9	Sequence 9, Appl
C 835	22	78.6	436	4	US-09-513-999C-11036	Sequence 11036, A	908	22	78.6	642	3	US-09-173-133-9	Sequence 9, Appl
C 836	22	78.6	437	4	US-09-489-039A-6132	Sequence 6132, Ap	909	22	78.6	642	4	US-09-580-236A-9	Sequence 9, Appl
C 837	22	78.6	438	4	US-09-248-796A-10519	Sequence 10519, A	910	22	78.6	654	4	US-09-270-767-30881	Sequence 30881, A
C 838	22	78.6	439	4	US-09-270-767-7198	Sequence 7198, Ap	911	22	78.6	658	3	US-08-961-083-27	Sequence 27, Appl
C 839	22	78.6	440	4	US-09-270-767-22480	Sequence 22480, A	912	22	78.6	658	3	US-09-536-784-27	Sequence 27, Appl
C 840	22	78.6	441	4	US-09-513-999C-11998	Sequence 11998, A	913	22	78.6	661	3	US-08-998-416-1100	Sequence 1100, Ap
C 841	22	78.6	442	4	US-09-621-976-13397	Sequence 13397, A	914	22	78.6	666	3	US-09-134-001C-416	Sequence 416, App
C 842	22	78.6	443	4	US-09-489-039A-6132	Sequence 6132, Ap	C 915	22	78.6	669	4	US-09-465-901-35	Sequence 35, Appl
C 843	22	78.6	444	4	US-08-956-171E-3176	Sequence 3176, Ap	916	22	78.6	678	4	US-09-328-352-2378	Sequence 2378, Ap
C 844	22	78.6	445	4	US-09-621-976-19188	Sequence 19188, A	917	22	78.6	684	4	US-09-248-796A-2807	Sequence 2807, Ap
C 845	22	78.6	446	4	US-08-781-986A-3176	Sequence 3176, Ap	C 918	22	78.6	690	3	US-08-348-548-5	Sequence 5, Appl
C 846	22	78.6	447	4	US-09-621-976-2058	Sequence 2058, Ap	C 919	22	78.6	690	3	US-08-252-991A-2533	Sequence 2533, Ap
C 847	22	78.6	448	4	US-09-513-999C-8540	Sequence 8540, Ap	C 920	22	78.6	690	5	PCT-US95-15716-5	Sequence 5, Appl
C 848	22	78.6	449	4	US-09-621-976-13397	Sequence 13397, A	C 921	22	78.6	693	4	US-09-270-767-25841	Sequence 25841, A
C 849	22	78.6	450	4	US-09-621-976-7931	Sequence 7931, Ap	922	22	78.6	696	4	US-09-583-110-2261	Sequence 2261, Ap
C 850	22	78.6	451	4	US-09-621-976-14205	Sequence 14205, A	923	22	78.6	697	3	US-08-998-416-958	Sequence 958, App
C 851	22	78.6	452	4	US-09-252-991A-8516	Sequence 8516, Ap	924	22	78.6	699	4	US-09-107-532A-2437	Sequence 2437, Ap
C 852	22	78.6	453	4	US-09-270-767-3772	Sequence 3772, Ap	925	22	78.6	699	4	US-08-248-796A-7245	Sequence 7245, Ap
C 853	22	78.6	454	4	US-09-270-767-19054	Sequence 19054, A	C 926	22	78.6	714	4	US-09-252-991A-6733	Sequence 6733, Ap
C 854	22	78.6	455	4	US-09-621-976-3610	Sequence 3610, Ap	927	22	78.6	717	4	US-09-328-352-2525	Sequence 2525, Ap
C 855	22	78.6	456	4	US-09-134-000C-1338	Sequence 1338, Ap	928	22	78.6	717	4	US-09-583-110-3625	Sequence 2625, Ap
C 856	22	78.6	457	4	US-09-543-681A-2528	Sequence 2528, Ap	C 929	22	78.6	726	4	US-09-328-475C-289	Sequence 289, App
C 857	22	78.6	458	4	US-08-867-890A-23	Sequence 23, Appl	C 930	22	78.6	731	4	US-09-976-594-191	Sequence 191, App
C 858	22	78.6	459	4	US-08-466-033-31	Sequence 31, Appl	931	22	78.6	732	4	US-08-107-532A-1011	Sequence 1011, Ap
C 859	22	78.6	460	4	US-08-464-733-31	Sequence 31, Appl	932	22	78.6	735	4	US-09-252-991A-6960	Sequence 6960, Ap
C 860	22	78.6	461	4	US-08-464-134-31	Sequence 31, Appl	C 933	22	78.6	739	3	US-09-385-982-159	Sequence 159, App
C 861	22	78.6	462	4	US-08-461-361-31	Sequence 31, Appl	C 934	22	78.6	743	4	US-09-328-475C-288	Sequence 288, App
C 862	22	78.6	463	4	US-08-461-361-31	Sequence 31, Appl	935	22	78.6	744	3	US-08-998-416-260	Sequence 260, App
C 863	22	78.6	464	4	US-08-485-910-31	Sequence 31, Appl	C 936	22	78.6	750	4	US-09-248-796A-4212	Sequence 4212, Ap
C 864	22	78.6	465	4	US-09-323-873A-7	Sequence 7, Appl	937	22	78.6	753	4	US-09-328-352-1602	Sequence 1602, Ap
C 865	22	78.6	466	4	US-10-027-983-14	Sequence 14, Appl	C 938	22	78.6	754	3	US-08-998-416-770	Sequence 770, App
C 866	22	78.6	467	4	US-08-920-440B-1	Sequence 1, Appl	C 939	22	78.6	755	4	US-09-712-016-23	Sequence 23, Appl
C 867	22	78.6	468	4	US-09-609-324A-1	Sequence 1, Appl	940	22	78.6	762	4	US-09-543-681A-1624	Sequence 1624, Ap
C 868	22	78.6	469	4	US-09-173-492-1	Sequence 1, Appl	C 941	22	78.6	764	2	US-08-935-886-9	Sequence 9, Appl
C 869	22	78.6	470	4	US-09-173-133-1	Sequence 1, Appl	942	22	78.6	765	4	US-09-252-991A-1409	Sequence 1409, Ap
C 870	22	78.6	471	4	US-09-163-533-1	Sequence 1, Appl	943	22	78.6	770	4	US-09-205-258-161	Sequence 161, App
C 871	22	78.6	472	4	US-09-580-236A-1	Sequence 1, Appl	C 944	22	78.6	778	4	US-09-270-767-10490	Sequence 10490, A
C 872	22	78.6	473	4	US-09-795-926-45	Sequence 45, Appl	C 945	22	78.6	780	3	US-08-447-402-6	Sequence 6, Appl
C 873	22	78.6	474	4	PCT-US95-12779-1	Sequence 1, Appl	946	22	78.6	780	3	US-08-838-151A-61	Sequence 61, Appl
C 874	22	78.6	475	4	US-09-346-408-9	Sequence 9, Appl	947	22	78.6	783	4	US-09-107-532A-1696	Sequence 1696, Ap
C 875	22	78.6	476	4	US-09-346-408-9	Sequence 9, Appl	C 948	22	78.6	792	4	US-09-244-111-3	Sequence 3, Appl
C 876	22	78.6	477	4	US-09-439-313-393	Sequence 393, App	C 949	22	78.6	793	4	US-09-270-767-13696	Sequence 13696, A
C 877	22	78.6	478	4	US-09-352-616A-393	Sequence 393, App	950	22	78.6	794	4	US-09-173-300-10	Sequence 10, Appl
C 878	22	78.6	479	4	US-09-636-215-393	Sequence 393, App	951	22	78.6	795	4	US-09-795-926-47	Sequence 47, Appl
C 879	22	78.6	480	4	US-09-685-166A-393	Sequence 393, App	C 952	22	78.6	799	4	US-09-270-767-9120	Sequence 9120, Ap
C 880	22	78.6	481	4	US-09-270-767-2593	Sequence 2593, Ap	C 953	22	78.6	799	4	US-09-270-767-24402	Sequence 24402, A
C 881	22	78.6	482	4	US-09-270-767-17875	Sequence 17875, A	954	22	78.6	801	4	US-09-252-991A-4545	Sequence 4545, Ap
C 882	22	78.6	483	4	US-09-679-426-393	Sequence 393, App	C 955	22	78.6	813	3	US-08-998-416-384	Sequence 384, App
C 883	22	78.6	484	4	US-09-146-950-3	Sequence 3, Appl	C 956	22	78.6	826	3	US-08-998-416-364	Sequence 364, App
C 884	22	78.6	485	4	US-09-107-532A-2613	Sequence 2613, Ap	C 957	22	78.6	831	3	US-08-998-416-328	Sequence 328, App
C 885	22	78.6	486	4	US-09-248-796A-13702	Sequence 13702, A	C 958	22	78.6	834	4	US-09-107-532A-3559	Sequence 3559, Ap
C 886	22	78.6	487	4	US-09-146-950-19	Sequence 19, Appl	C 959	22	78.6	840	4	US-09-270-767-9170	Sequence 9170, Ap
C 887	22	78.6	488	4	US-09-614-221A-304	Sequence 304, App	C 960	22	78.6	840	4	US-09-270-767-24452	Sequence 24452, A
C 888	22	78.6	489	4	US-09-328-111-155	Sequence 155, App	961	22	78.6	849	4	US-09-252-991A-10108	Sequence 10108, A

RESULT 3

US-09-107-532A-1525
; Sequence 1525, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bueh
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1525:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1065 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...1065
; SEQUENCE DESCRIPTION: SEQ ID NO: 1525:
US-09-107-532A-1525

Alignment Scores:
Pred. No.: 218 Length: 1065
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-107-532A-1525 (1-1065)

Qy 1 GlyTyr***ValGluGlu 6
Db 361 GGTACACGGTTGAGAA 378

RESULT 4

US-08-907-166-9
; Sequence 9, Application US/08907166
; Patent No. 5948666
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter

; APPLICANT: Mather, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/08/907,166
; CURRENT FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Desulfurococcus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2286)
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1801)..(1801)
; OTHER INFORMATION: s at position 1801 is either c or g
US-08-907-166-9

Alignment Scores:
Pred. No.: 517 Length: 2289
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-5 (1-6) x US-08-907-166-9 (1-2289)

Qy 1 GlyTyr***ValGluGlu 6
Db 2065 GGTACACGGTGGAGGAG 2082

RESULT 5

US-09-391-340-9
; Sequence 9, Application US/09391340A
; Patent No. 6492511
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Mather, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/09/391,340A
; CURRENT FILING DATE: 1999-09-07
; EARLIER APPLICATION NUMBER: US 08/907,166
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Desulfurococcus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2286)
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1801)..(1801)
; OTHER INFORMATION: s at position 1801 is either c or g
US-09-391-340-9

Alignment Scores:
Pred. No.: 517 Length: 2289
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-391-340-9 (1-2289)

Qy 1 GlyTyr***ValGluGlu 6

```
Db 2065 GGGTACACGGTGGAGGAG 2082
||||| |||||||
RESULT 6
US-09-252-991A-11271/c
; Sequence 11271, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11271
; LENGTH: 2547
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11271
Alignment Scores:
Pred. No.: 584 Length: 2547
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 4 Gaps: 0
US-10-030-194A-5 (1-6) x US-09-252-991A-11271 (1-2547)
Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 42 GGGTACACGGTGGAGGAA 25
RESULT 7
US-09-252-991A-10912
; Sequence 10912, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10912
; LENGTH: 2964
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10912
Alignment Scores:
Pred. No.: 693 Length: 2964
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 4 Gaps: 0
US-10-030-194A-5 (1-6) x US-09-252-991A-10912 (1-2964)
Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 2065 GGGTACACGGTGGAGGAG 2082
||||| |||||||
RESULT 8
US-08-956-171E-4/c
; Sequence 4, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13321 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-956-171E-4
Alignment Scores:
Pred. No.: 3.8e+03 Length: 13321
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 4 Gaps: 0
US-10-030-194A-5 (1-6) x US-08-956-171E-4 (1-13321)
Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 7899 GGATACCTCTGTAGAAGAA 7882
RESULT 9
US-08-781-986A-4/c
; Sequence 4, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
```



```

RESULT 12
US-09-513-999C-36409
; Sequence 36409, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 36409
; LENGTH: 107
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 63
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 97
; OTHER INFORMATION: w=a or t
US-09-513-999C-36409

Alignment Scores:
Pred. No.: 30.5 Length: 107
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-513-999C-36409 (1-107)

Qy 1 GlyTyr***ValGluGlu 6
Db 27 GGGTATCTGTTGGAAGAA 44

RESULT 13
US-08-313-050-9
; Sequence 9, Application US/08313050
; Patent No. 5585256
; GENERAL INFORMATION:
; APPLICANT: Dorreich, Kurt
; APPLICANT: Christensen, Flemming M.
; APPLICANT: Schnell, Yvette
; APPLICANT: Mischler, Marcel
; APPLICANT: Dalboge, Henrik
; APPLICANT: Heldt-Hansen, Hans P.
; TITLE OF INVENTION: A NOVEL ENZYME AND A DNA SEQUENCE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5585256 No. 5585256disk of No. 5585256th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313.050

```

```

; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0420/92
; FILING DATE: 27-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK PCT/DK93/00109
; FILING DATE: 29-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 3730.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 159 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-313-050-9

Alignment Scores:
Pred. No.: 47.7 Length: 159
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 1 Gaps: 0

US-10-030-194A-5 (1-6) x US-08-313-050-9 (1-159)

Qy 1 GlyTyr***ValGluGlu 6
Db 121 GGGTACGTCGTTGAAGAG 138

RESULT 14
US-09-513-999C-33426
; Sequence 33426, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 33426
; LENGTH: 170
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-33426

Alignment Scores:
Pred. No.: 51.4 Length: 170
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-513-999C-33426 (1-170)

Qy 1 GlyTyr***ValGluGlu 6

```

Db 57 GGGTATAGGTGGAGAA 74

RESULT 15

US-09-621-976-9290

; Sequence 9290, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 9290

; LENGTH: 239

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-621-976-9290

Alignment Scores:

Pred. No.:	75.6	Length:	239
Score:	26.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	92.86%	Indels:	0
DB:	4	Gaps:	0

US-10-030-194A-5 (1-6) x US-09-621-976-9290 (1-239)

QY 1 GlyTyr***ValGluGlu 6

Db 23 GGGTATCTGGTGGAGAA 40

RESULT 16

US-09-513-999C-17433

; Sequence 17433, Application US/09513999C

; Patent No. 6783961

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclert, A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; FILE REFERENCE: 59 US2.REG

; CURRENT APPLICATION NUMBER: US/09/513,999C

; CURRENT FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/122,487

; PRIOR FILING DATE: 1999-02-26

; NUMBER OF SEQ ID NOS: 36681

; SOFTWARE: Patent.pm

; SEQ ID NO 17433

; LENGTH: 262

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc_feature

; LOCATION: 211_feature

; OTHER INFORMATION: w=a or t

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 243

; OTHER INFORMATION: w=a or t

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 254

; OTHER INFORMATION: k=g or t

US-09-513-999C-17433

Alignment Scores:

Pred. No.:	83.9	Length:	262
Score:	83.9	Matches:	262
Percent Similarity:	83.9%	Conservative:	0
Best Local Similarity:	83.9%	Mismatches:	0
Query Match:	92.86%	Indels:	0
DB:	4	Gaps:	0

Score: 26.00 Matches: 5

Percent Similarity: 83.33% Conservative: 0

Best Local Similarity: 83.33% Mismatches: 1

Query Match: 92.86% Indels: 0

DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-513-999C-17433 (1-262)

QY 1 GlyTyr***ValGluGlu 6

Db 93 GGTATAAGTAGAAG 110

RESULT 17

US-09-513-999C-3343/c

; Sequence 3343, Application US/09513999C

; Patent No. 6783961

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclert, A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; FILE REFERENCE: 59 US2.REG

; CURRENT APPLICATION NUMBER: US/09/513,999C

; CURRENT FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/122,487

; PRIOR FILING DATE: 1999-02-26

; NUMBER OF SEQ ID NOS: 36681

; SOFTWARE: Patent.pm

; SEQ ID NO 3343

; LENGTH: 263

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: CDS

; LOCATION: 84_263

US-09-513-999C-3343

Alignment Scores:

Pred. No.:	84.2	Length:	263
Score:	26.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	92.86%	Indels:	0
DB:	4	Gaps:	0

US-10-030-194A-5 (1-6) x US-09-513-999C-3343 (1-263)

QY 1 GlyTyr***ValGluGlu 6

Db 31 GGGTATTGGTGAAGAA 14

RESULT 18

US-09-134-001C-1648/c

; Sequence 1648, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 1648

; LENGTH: 270

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-1648

```

Query Match:      92.86%      Indels:      0
DB:              1          Gaps:      0

US-10-030-194A-5 (1-6) x US-08-313-050-17 (1-276)

QY      1  GvTyr***valgluglu 6
          ||||| |||||
Db      121 GGTACGTCGTTGAAGAG 138

RESULT 20
US-08-808-982-4
; Sequence 4, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsey
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-808-982-4

Alignment Scores:
Pred. No.:      99.6      Length:      305
Score:          26.00     Matches:      5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match:    92.86%   Indels:      0
DB:             2       Gaps:      0

US-10-030-194A-5 (1-6) x US-08-808-982-4 (1-305)

QY      1  GvTyr***valgluglu 6
          ||||| |||||
Db      249 GGATACCTGTCGAGGAG 266

RESULT 21
US-09-306-902A-4
; Sequence 4, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc

```

```
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA: US/09/306,902A
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
;
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-306-902A-4

Alignment Scores:
Pred. No.: 99.6 Length: 305
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-306-902A-4 (1-305)
QY 1 GlyTyr***ValGluGlu 6
Db 249 GGATACCTGGTGGAGGAG 266

RESULT 22
US-09-513-999C-22044/c
; Sequence 22044, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2 REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 22044
; LENGTH: 305
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-513-999C-22044

Alignment Scores:
Pred. No.: 99.6 Length: 305
Score: 26.00 Matches: 4
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 66.67% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-513-999C-22044 (1-305)
QY 1 GlyTyr***ValGluGlu 6
Db 83 GGTACAGCATAGAGAA 66

RESULT 23
US-09-107-532A-2175
; Sequence 2175, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Denise
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2175:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 324 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...324
; SEQUENCE DESCRIPTION: SEQ ID NO: 2175:
US-09-107-532A-2175

Alignment Scores:
Pred. No.: 107 Length: 324
Score: 26.00 Matches: 4
```

Percent Similarity: 83.33% Conservative: 1
 Best Local Similarity: 66.67% Mismatches: 1
 Query Match: 92.86% Indels: 0
 DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-107-532A-2175 (1-324)

QY 1 GlyTyr***ValGluGlu 6
 ||||| :|||||
 Db 256 GGATATAGTATTGAGAG 273

RESULT 24
 US-09-513-999C-35661/c
 ; Sequence 35661, Application US/09513999C
 ; Patent No. 6783961
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Duclert, A.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
 ; Patent No. 6783961
 ; FILE REFERENCE: 59.US2.REG
 ; CURRENT APPLICATION NUMBER: US/09/513.999C
 ; CURRENT FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/122,487
 ; PRIOR FILING DATE: 1999-02-26
 ; NUMBER OF SEQ ID NOS: 36681
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 35661
 ; LENGTH: 383
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-513-999C-35661

Alignment Scores:
 Pred. No.: 129 Length: 383
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 92.86% Indels: 0
 DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-513-999C-35661 (1-383)

QY 1 GlyTyr***ValGluGlu 6
 ||||| :|||||
 Db 150 GGCTATAAAGTCGAAGAA 133

RESULT 25
 US-09-134-000C-165
 ; Sequence 165, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; PRIOR FILING DATE: 1997-08-15
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 165
 ; LENGTH: 387
 ; TYPE: DNA
 ; ORGANISM: Enterococcus faecalis
 US-09-134-000C-165

Alignment Scores:
 Pred. No.: 130 Length: 387
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0

Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 92.86% Indels: 0
 DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-134-000C-165 (1-387)

QY 1 GlyTyr***ValGluGlu 6
 ||||| :|||||
 Db 274 GGCTATCGGTGAAGAA 291

RESULT 26
 US-09-513-999C-15816
 ; Sequence 15816, Application US/09513999C
 ; Patent No. 6783961
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Duclert, A.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
 ; Patent No. 6783961
 ; FILE REFERENCE: 59.US2.REG
 ; CURRENT APPLICATION NUMBER: US/09/513.999C
 ; CURRENT FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/122,487
 ; PRIOR FILING DATE: 1999-02-26
 ; NUMBER OF SEQ ID NOS: 36681
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 15816
 ; LENGTH: 403
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc_feature
 ; LOCATION: 389
 ; OTHER INFORMATION: r-a or g
 US-09-513-999C-15816

Alignment Scores:
 Pred. No.: 137 Length: 403
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 92.86% Indels: 0
 DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-513-999C-15816 (1-403)

QY 1 GlyTyr***ValGluGlu 6
 ||||| :|||||
 Db 289 GGCTACAATGTTGAGGAG 306

RESULT 27
 US-09-513-999C-28426
 ; Sequence 28426, Application US/09513999C
 ; Patent No. 6783961
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Duclert, A.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
 ; Patent No. 6783961
 ; FILE REFERENCE: 59.US2.REG
 ; CURRENT APPLICATION NUMBER: US/09/513.999C
 ; CURRENT FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/122,487
 ; PRIOR FILING DATE: 1999-02-26
 ; NUMBER OF SEQ ID NOS: 36681
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 28426
 ; LENGTH: 404
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-513-999C-28426


```
DB:                                     3          Gaps:          0
US-10-030-194a-5 (1-6) x US-09-280-116-153 (1-495)

Qy      1 GlyTyr***ValGluGlu 6
      ||||| ||||| |||||
Db      57 GCCTATGAGTGAAGAG 74

RESULT 31
US-09-205-258-91/c
; Sequence 91, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-205-258-91

Alignment Scores:
Pred. No.:      204      Length:      575
Score:          26.00    Matches:      5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match:      92.86% Indels:      0
DB:               4      Gaps:      0

US-10-030-194a-5 (1-6) x US-09-205-258-91 (1-575)

Qy      1 GlyTyr***ValGluGlu 6
      ||||| ||||| |||||
Db      551 GGCTACCTAGTAGAGGAA 534

RESULT 32
US-09-585-173B-19
; Sequence 19, Application US/09585173B
; Patent No. 6570063
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Gutteridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelataase
; FILE REFERENCE: Bb1370 US NA
; CURRENT APPLICATION NUMBER: US/09/585,173B
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (397)
; OTHER INFORMATION: n is a, c, g or t
; FEATURE:
```



```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-253-155A-17

Alignment Scores:
Pred. No.: 285 Length: 771
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 1 Gaps: 0

US-10-030-194A-5 (1-6) x US-08-253-155A-17 (1-771)

Qy 1 GlyTyr***ValGluGlu 6
Db 325 GGATACCTGGTGGAGGAG 308

RESULT 35
US-09-585-173B-21
; Sequence 21, Application US/09585173B
; Patent No. 6570063
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Gutteridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelatase
; FILE REFERENCE: BB1370 US NA
; CURRENT APPLICATION NUMBER: US/09/585.173B
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 782
; TYPE: DNA
; ORGANISM: Zea mays
US-09-585-173B-21

Alignment Scores:
Pred. No.: 289 Length: 782
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-585-173B-21 (1-782)

Qy 1 GlyTyr***ValGluGlu 6
Db 78 GGCTACCTCGTCGAGGAG 95

RESULT 36
US-08-313-050-10
; Sequence 10, Application US/08313050
; Patent No. 5585256
; GENERAL INFORMATION:
; APPLICANT: Dorreich, Kurt
; APPLICANT: Christensen, Flemming M.
; APPLICANT: Schnell, Yvette
; APPLICANT: Mischler, Marcel
; APPLICANT: Dalboge, Henrik
; APPLICANT: Heldt-Hansen, Hans P.
; TITLE OF INVENTION: A NOVEL ENZYME AND A DNA SEQUENCE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 55852560 No. 5585256disk of No. 5585256th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313.050
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0420/92
; FILING DATE: 27-MAR-1992
; PRIOR APPLICATION NUMBER: DK PCT/DK93/00109
; FILING DATE: 29-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 3730.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-313-050-10

Alignment Scores:
Pred. No.: 354 Length: 935
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 1 Gaps: 0

US-10-030-194A-5 (1-6) x US-08-313-050-10 (1-935)

Qy 1 GlyTyr***ValGluGlu 6
Db 733 GGGTACGTCGTTGAAGAG 750

RESULT 37
US-09-134-000C-86
; Sequence 86, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
```


Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-170-097-204 (1-1001)
Qy 1 GlyTyr***ValGluGlu 6
Db 261 GGTACAGGTTGAGGAG 278
RESULT 40
US-09-134-001C-1568
; Sequence 1568, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1568
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1568
Alignment Scores:
Pred. No.: 414 Length: 1074
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Indels: 0
Query Match: 92.86%
DB: 3 Gaps: 0
US-10-030-194A-5 (1-6) x US-09-134-001C-1568 (1-1074)
Qy 1 GlyTyr***ValGluGlu 6
Db 961 GGTATGTTGTTGAAGAA 978
RESULT 41
US-09-170-279-17
; Sequence 17, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU34800S
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-17
Alignment Scores:
Pred. No.: 415 Length: 1077
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1

Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0
US-10-030-194A-5 (1-6) x US-09-710-279-17 (1-1077)
Qy 1 GlyTyr***ValGluGlu 6
Db 814 GGTATGAAGTTGAAGAG 831
RESULT 42
US-09-173-300-14
; Sequence 14, Application US/09173300
; Patent No. 6451581
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1126
; CURRENT APPLICATION NUMBER: US/09/173,300
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 60/063,423
; EARLIER FILING DATE: 1997 October 28
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 14
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-173-300-14
Alignment Scores:
Pred. No.: 419 Length: 1086
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Indels: 0
Query Match: 92.86%
DB: 4 Gaps: 0
US-10-030-194A-5 (1-6) x US-09-173-300-14 (1-1086)
Qy 1 GlyTyr***ValGluGlu 6
Db 692 GCCTATCAGTTGAAGAA 709
RESULT 43
US-09-134-001C-2566
; Sequence 2566, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2566
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2566
Alignment Scores:
Pred. No.: 421 Length: 1089
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0

Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-134-001C-2566 (1-1089)

QY 1 GlyTyr***ValGluGlu 6
DB 826 GGTATGAAGTTGAAGAG 843

RESULT 44

US-08-183-214-11
; Sequence 11, Application US/08183214
; Patent No. 5716816
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Stanley, Sally J.
; APPLICANT: Nightingale, Maria S.
; APPLICANT: Murtagh, Jr., James J.
; APPLICANT: Monaco, Lucia
; APPLICANT: Takada, Tatsuyuki
; TITLE OF INVENTION: CLONES ENCODING MAMMALIAN
; ADP-RIBOSYLARGININE HYDROLASES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/183,214
; FILING DATE: 14-JAN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,231
; FILING DATE: 22-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 11:
; LENGTH: 1109 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1038
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 338..339
; OTHER INFORMATION: /note= "Nucleotide at position 338
; OTHER INFORMATION: may be either G or A, and the amino acid encoded
; OTHER INFORMATION: by this codon is either serine or asparagine."

US-08-183-214-11
Alignment Scores:
Pred. No.: 429 Length: 1109
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1

Query Match: 92.86% Indels: 0
DB: 1 Gaps: 0

US-10-030-194A-5 (1-6) x US-08-183-214-11 (1-1109)

QY 1 GlyTyr***ValGluGlu 6
DB 595 GGCTACTTTGTAGAGAA 612

RESULT 45

US-09-543-681A-3142
; Sequence 3142, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO: 3142
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3142

Alignment Scores:
Pred. No.: 430 Length: 1110
Score: 26.00 Matches: 4
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 66.67% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-543-681A-3142 (1-1110)

QY 1 GlyTyr***ValGluGlu 6
DB 871 GGATATCTATTGAGAA 888

RESULT 46

US-09-252-991A-8501/c
; Sequence 8501, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 8501
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8501

Alignment Scores:
Pred. No.: 438 Length: 1128
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Tomato Yellow Leaf Curl Virus
; STRAIN: Israel
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 37..1107
; PUBLICATION INFORMATION:
; AUTHORS: Navot, N
; AUTHORS: Pichersky, R
; AUTHORS: Zeidan, D
; AUTHORS: Zamir, D
; AUTHORS: Czosnek, H
; TITLE: Tomato yellow leaf curl virus: A
; TITLE: whitefly-transmitted geminivirus with a single
; TITLE: genomic component.
; JOURNAL: Virology
; VOLUME: 185
; PAGES: 151-168
; DATE: 1991
US-08-838-151A-19
Alignment Scores:
Pred. No.: 445 Length: 1145
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0
US-10-030-194A-5 (1-6) x US-08-838-151A-19 (1-1145)
Qy 1 GlyTyr***ValGluGlu 6
Db 270 GGATACCAAGTCGAAGAA 253
RESULT 50
US-08-838-151A-23/c
; Sequence 23, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
; APPLICANT: Stout, John T
; APPLICANT: Luu, Hang T
; APPLICANT: Maxwell, Douglas
; APPLICANT: Ahlquist, Paul
; APPLICANT: Hanson, Steve
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,151A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueller, Lisa V
; REGISTRATION NUMBER: 38,978
; REFERENCE/DOCKET NUMBER: SVS3801P0260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400

; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Tomato Yellow Leaf Curl Geminivirus
; STRAIN: Israel
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 37..1107
US-08-838-151A-23
Alignment Scores:
Pred. No.: 445 Length: 1145
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0
US-10-030-194A-5 (1-6) x US-08-838-151A-23 (1-1145)
Qy 1 GlyTyr***ValGluGlu 6
Db 270 GGATACCAAGTCGAAGAA 253
RESULT 51
US-08-838-151A-26/c
; Sequence 26, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
; APPLICANT: Stout, John T
; APPLICANT: Luu, Hang T
; APPLICANT: Maxwell, Douglas
; APPLICANT: Ahlquist, Paul
; APPLICANT: Hanson, Steve
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,151A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueller, Lisa V
; REGISTRATION NUMBER: 38,978
; REFERENCE/DOCKET NUMBER: SVS3801P0260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```
/
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Tomato Yellow Leaf Curl Virus
/ STRAIN: Israel
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 37..1107
/ PUBLICATION INFORMATION:
/ AUTHORS: Navot, N
/ AUTHORS: Pichersky, R
/ AUTHORS: Zeidan, D
/ AUTHORS: Zamir, D
/ AUTHORS: Czosnek, H
/ TITLE: Tomato yellow leaf curl virus: A
/ TITLE: whitefly-transmitted geminivirus with a single
/ TITLE: genomic component.
/ JOURNAL: Virology
/ VOLUME: 185
/ PAGES: 151-168
/ DATE: 1991
/
US-08-838-151A-26

Alignment Scores:
Pred. No.: 445 Length: 1145
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-08-838-151A-26 (1-1145)

Oy 1 GlyTyr***ValGluGlu 6
Db 270 GGATACCAAGTCGAAGAA 253

RESULT 52
US-08-838-151A-29/c
; Sequence 29, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
; APPLICANT: Stout, John T
; APPLICANT: Luu, Hang T
; APPLICANT: Maxwell, Douglas
; APPLICANT: Ahlquist, Paul
; APPLICANT: Hanson, Steve
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus
; TITLE OF INVENTION: Genes
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838.151A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueller, Lisa V
; REGISTRATION NUMBER: 38,978
; REFERENCE/DOCKET NUMBER: SVS3801P0260
```

```
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312-616-5400
/ TELEFAX: 312-616-5460
/ INFORMATION FOR SEQ ID NO: 29:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1145 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Tomato Yellow Leaf Curl Virus
/ STRAIN: Israel
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 37..1107
/ PUBLICATION INFORMATION:
/ AUTHORS: Navot, N
/ AUTHORS: Pichersky, R
/ AUTHORS: Zeidan, D
/ AUTHORS: Zamir, D
/ AUTHORS: Czosnek, H
/ TITLE: Tomato yellow leaf curl virus: A
/ TITLE: whitefly-transmitted geminivirus with a single
/ TITLE: genomic component.
/ JOURNAL: Virology
/ VOLUME: 185
/ PAGES: 151-168
/ DATE: 1991
/
US-08-838-151A-29

Alignment Scores:
Pred. No.: 445 Length: 1145
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-08-838-151A-29 (1-1145)

Oy 1 GlyTyr***ValGluGlu 6
Db 270 GGATACCAAGTCGAAGAA 253

RESULT 53
US-08-809-103B-1/c
; Sequence 1, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GROENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809.103B
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94.11040
; FILING DATE: 15-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/01192
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1148 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1077
; US-08-809-103B-1

Alignment Scores: 446 1148
Pred. No.: 26.00 Matches: 5
Score: 83.33% Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 92.86%
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-08-809-103B-1 (1-1148)

Qy 1 GlyTyr***ValGluGlu 6
Db 240 GGATACCGAGTCGAAGAA 223

RESULT 54

US-08-809-103B-3/c
; Sequence 3, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GRONENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,103B
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94.11040
; FILING DATE: 15-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/01192
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:

; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1077
; US-08-809-103B-3

Alignment Scores: 447 1150
Pred. No.: 26.00 Matches: 5
Score: 83.33% Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 92.86%
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-08-809-103B-3 (1-1150)

Qy 1 GlyTyr***ValGluGlu 6
Db 240 GGATACCGAGTCGAAGAA 223

RESULT 55

US-08-809-103B-5/c
; Sequence 5, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GRONENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,103B
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94.11040
; FILING DATE: 15-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/01192
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:


```
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2485
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2485

Alignment Scores:
Pred. No.: 551 Length: 1383
Score: 26.00 Matches: 4
Percent Similarity: 83.33% Conservative: 1
Best Local Similarity: 66.67% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-134-001C-2485 (1-1383)

Qy 1 GlyTyr***ValGluGlu 6
| | | | | : | | | | |
Db 70 GGTATTCTATTGAAGAA 87

RESULT 60
US-08-872-302-1/c
; Sequence 1, Application US/08872302
; Patent No. 5846784
; GENERAL INFORMATION:
; APPLICANT: Hitz, William D
; TITLE OF INVENTION: Fatty Acid Modifying Enzymes From
; NUMBER OF SEQUENCES: 10 Developing Seeds of Vernonia galamensis
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. duPont de Nemours and Co.
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,302
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Majarian, William R
; REGISTRATION NUMBER: P-41,173
; REFERENCE/DOCKET NUMBER: BB-1084
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4926
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 134..1279
US-08-872-302-1

Alignment Scores:
Pred. No.: 593 Length: 1476
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 2 Gaps: 0
```

```
US-10-030-194A-5 (1-6) x US-08-872-302-1 (1-1476)

Qy 1 GlyTyr***ValGluGlu 6
| | | | | : | | | | |
Db 340 GGCTACATAGTAGAGGAG 323

RESULT 61
US-09-117-853-7
; Sequence 7, Application US/09117853
; Patent No. 6307126
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; CURRENT FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-117-853-7

Alignment Scores:
Pred. No.: 667 Length: 1636
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-117-853-7 (1-1636)

Qy 1 GlyTyr***ValGluGlu 6
| | | | | : | | | | |
Db 1596 GGTATCGGTGAGGAG 1613

RESULT 62
US-09-911-154-7
; Sequence 7, Application US/09911154
; Patent No. 6478809
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,154
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-154-7

Alignment Scores:
```

Pred. No.: 667 Length: 1636
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-911-154-7 (1-1636)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 1596 GGTATCGGTGGAGGAG 1613

RESULT 63

US-09-911-514-7
; Sequence 7, Application US/09911514
; Patent No. 6794560
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,514
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-514-7

Alignment Scores:
Pred. No.: 667 Length: 1636
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-911-514-7 (1-1636)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 1596 GGTATCGGTGGAGGAG 1613

RESULT 64

US-09-117-853-5
; Sequence 5, Application US/09117853
; Patent No. 6307126
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; CURRENT FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-117-853-5

Alignment Scores:
Pred. No.: 669 Length: 1642
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-117-853-5 (1-1642)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 1602 GGTATCGGTGGAGGAG 1619

RESULT 65

US-09-117-853-9
; Sequence 9, Application US/09117853
; Patent No. 6307126
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; CURRENT FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-117-853-9

Alignment Scores:
Pred. No.: 669 Length: 1642
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-117-853-9 (1-1642)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 1602 GGTATCGGTGGAGGAG 1619

RESULT 66

US-09-911-154-5
; Sequence 5, Application US/09911154
; Patent No. 6478809
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,154
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853

QY	DB	Length:	Matches:	Score:	Percent Similarity:	Best Local Similarity:	Mismatches:	Indels:	Gaps:
QY	1 GlyTyr***ValGluGlu 6	1642	5	669	26.00	83.33%	0	0	0
DB	1602 GGTATCGGTGGAGGAG 1619								
<p>US-10-030-194A-5 (1-6) x US-09-911-154-5 (1-1642)</p> <p>Alignment Scores:</p> <p>Pred. No.: 669</p> <p>Score: 26.00</p> <p>Percent Similarity: 83.33%</p> <p>Best Local Similarity: 83.33%</p> <p>Mismatches: 0</p> <p>Indels: 0</p> <p>Query Match: 92.86%</p> <p>DB: 4</p>									
<p>US-09-911-154-9</p> <p>Sequence 9, Application US/09911154</p> <p>Patent No. 6478809</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Harberd, Nicholas P</p> <p>APPLICANT: Peng, Jinrong</p> <p>APPLICANT: Carol, Pierre</p> <p>TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana</p> <p>FILE REFERENCE: 620-158</p> <p>CURRENT APPLICATION NUMBER: US/09/911,154</p> <p>PRIOR FILING DATE: 2001-07-25</p> <p>PRIOR APPLICATION NUMBER: PCT/GB97/00390</p> <p>PRIOR FILING DATE: 1997-02-12</p> <p>PRIOR APPLICATION NUMBER: GB 9602796.6</p> <p>NUMBER OF SEQ ID NOS: 12</p> <p>SOFTWARE: PatentIn Ver. 2.0</p> <p>SEQ ID NO 5</p> <p>LENGTH: 1642</p> <p>TYPE: DNA</p> <p>ORGANISM: Arabidopsis thaliana</p> <p>US-09-911-154-5</p>									
<p>RESULT 67</p> <p>US-09-911-154-9</p> <p>Sequence 9, Application US/09911154</p> <p>Patent No. 6478809</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Harberd, Nicholas P</p> <p>APPLICANT: Peng, Jinrong</p> <p>APPLICANT: Carol, Pierre</p> <p>APPLICANT: Richards, Donald E</p> <p>TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana</p> <p>FILE REFERENCE: 620-158</p> <p>CURRENT APPLICATION NUMBER: US/09/911,154</p> <p>PRIOR FILING DATE: 2001-07-25</p> <p>PRIOR APPLICATION NUMBER: US 09/117,853</p> <p>PRIOR FILING DATE: 1998-08-12</p> <p>PRIOR APPLICATION NUMBER: PCT/GB97/00390</p> <p>PRIOR FILING DATE: 1997-02-12</p> <p>PRIOR APPLICATION NUMBER: GB 9602796.6</p> <p>NUMBER OF SEQ ID NOS: 12</p> <p>SOFTWARE: PatentIn Ver. 2.0</p> <p>SEQ ID NO 9</p> <p>LENGTH: 1642</p> <p>TYPE: DNA</p> <p>ORGANISM: Arabidopsis thaliana</p> <p>US-09-911-154-9</p>									
<p>Alignment Scores:</p> <p>Pred. No.: 669</p> <p>Score: 26.00</p> <p>Percent Similarity: 83.33%</p> <p>Best Local Similarity: 83.33%</p> <p>Mismatches: 0</p> <p>Indels: 0</p> <p>Query Match: 92.86%</p> <p>DB: 4</p>									
<p>US-10-030-194A-5 (1-6) x US-09-911-154-9 (1-1642)</p> <p>Alignment Scores:</p> <p>Pred. No.: 669</p> <p>Score: 26.00</p> <p>Percent Similarity: 83.33%</p> <p>Best Local Similarity: 83.33%</p> <p>Mismatches: 0</p> <p>Indels: 0</p> <p>Query Match: 92.86%</p> <p>DB: 4</p>									
<p>US-09-911-154-9</p> <p>Sequence 9, Application US/09911154</p> <p>Patent No. 6478809</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Harberd, Nicholas P</p> <p>APPLICANT: Peng, Jinrong</p> <p>APPLICANT: Carol, Pierre</p> <p>APPLICANT: Richards, Donald E</p> <p>TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana</p> <p>FILE REFERENCE: 620-158</p> <p>CURRENT APPLICATION NUMBER: US/09/911,154</p> <p>PRIOR FILING DATE: 2001-07-25</p> <p>PRIOR APPLICATION NUMBER: US 09/117,853</p> <p>PRIOR FILING DATE: 1998-08-12</p> <p>PRIOR APPLICATION NUMBER: PCT/GB97/00390</p> <p>PRIOR FILING DATE: 1997-02-12</p> <p>PRIOR APPLICATION NUMBER: GB 9602796.6</p> <p>NUMBER OF SEQ ID NOS: 12</p> <p>SOFTWARE: PatentIn Ver. 2.0</p> <p>SEQ ID NO 9</p> <p>LENGTH: 1642</p> <p>TYPE: DNA</p> <p>ORGANISM: Arabidopsis thaliana</p> <p>US-09-911-154-9</p>									
<p>Alignment Scores:</p> <p>Pred. No.: 669</p> <p>Score: 26.00</p> <p>Percent Similarity: 83.33%</p> <p>Best Local Similarity: 83.33%</p> <p>Mismatches: 0</p> <p>Indels: 0</p> <p>Query Match: 92.86%</p> <p>DB: 4</p>									
<p>US-10-030-194A-5 (1-6) x US-09-911-154-9 (1-1642)</p> <p>Alignment Scores:</p> <p>Pred. No.: 669</p> <p>Score: 26.00</p> <p>Percent Similarity: 83.33%</p> <p>Best Local Similarity: 83.33%</p> <p>Mismatches: 0</p> <p>Indels: 0</p> <p>Query Match: 92.86%</p> <p>DB: 4</p>									
<p>US-09-911-154-9</p> <p>Sequence 9, Application US/09911154</p> <p>Patent No. 6478809</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Harberd, Nicholas P</p> <p>APPLICANT: Peng, Jinrong</p> <p>APPLICANT: Carol, Pierre</p> <p>APPLICANT: Richards, Donald E</p> <p>TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana</p> <p>FILE REFERENCE: 620-158</p> <p>CURRENT APPLICATION NUMBER: US/09/911,154</p> <p>PRIOR FILING DATE: 2001-07-25</p> <p>PRIOR APPLICATION NUMBER: US 09/117,853</p> <p>PRIOR FILING DATE: 1998-08-12</p> <p>PRIOR APPLICATION NUMBER: PCT/GB97/00390</p> <p>PRIOR FILING DATE: 1997-02-12</p> <p>PRIOR APPLICATION NUMBER: GB 9602796.6</p> <p>NUMBER OF SEQ ID NOS: 12</p> <p>SOFTWARE: PatentIn Ver. 2.0</p> <p>SEQ ID NO 9</p> <p>LENGTH: 1642</p> <p>TYPE: DNA</p> <p>ORGANISM: Arabidopsis thaliana</p> <p>US-09-911-154-9</p>									
<p>Alignment Scores:</p> <p>Pred. No.: 669</p> <p>Score: 26.00</p> <p>Percent Similarity: 83.33%</p> <p>Best Local Similarity: 83.33%</p> <p>Mismatches: 0</p> <p>Indels: 0</p> <p>Query Match: 92.86%</p> <p>DB: 4</p>									
<p>US-10-030-194A-5 (1-6) x US-09-911-154-9 (1-1642)</p> <p>Alignment Scores:</p> <p>Pred. No.: 669</p> <p>Score: 26.00</p> <p>Percent Similarity: 83.33%</p> <p>Best Local Similarity: 83.33%</p> <p>Mismatches: 0</p> <p>Indels: 0</p> <p>Query Match: 92.86%</p> <p>DB: 4</p>									
<p>US-09-911-154-9</p> <p>Sequence 9, Application US/09911154</p> <p>Patent No. 6478809</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Harberd, Nicholas P</p> <p>APPLICANT: Peng, Jinrong</p> <p>APPLICANT: Carol, Pierre</p> <p>APPLICANT: Richards, Donald E</p> <p>TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana</p> <p>FILE REFERENCE: 620-158</p> <p>CURRENT APPLICATION NUMBER: US/09/911,154</p> <p>PRIOR FILING DATE: 2001-07-25</p> <p>PRIOR APPLICATION NUMBER: US 09/117,853</p> <p>PRIOR FILING DATE: 1998-08-12</p> <p>PRIOR APPLICATION NUMBER: PCT/GB97/00390</p> <p>PRIOR FILING DATE: 1997-02-12</p> <p>PRIOR APPLICATION NUMBER: GB 9602796.6</p> <p>NUMBER OF SEQ ID NOS: 12</p> <p>SOFTWARE: PatentIn Ver. 2.0</p> <p>SEQ ID NO 9</p> <p>LENGTH: 1642</p> <p>TYPE: DNA</p> <p>ORGANISM: Arabidopsis thaliana</p> <p>US-09-911-154-9</p>									
<p>Alignment Scores:</p> <p>Pred. No.: 669</p> <p>Score: 26.00</p> <p>Percent Similarity: 83.33%</p> <p>Best Local Similarity: 83.33%</p> <p>Mismatches: 0</p> <p>Indels: 0</p> <p>Query Match: 92.86%</p> <p>DB: 4</p>									
<p>US-10-030-194A-5 (1-6) x US-09-911-154-9 (1-1642)</p> <p>Alignment Scores:</p> <p>Pred. No.: 669</p> <p>Score: 26.00</p> <p>Percent Similarity: 83.33%</p> <p>Best Local Similarity: 83.33%</p> <p>Mismatches: 0</p> <p>Indels: 0</p> <p>Query Match: 92.86%</p> <p>DB: 4</p>									
<p>US-09-911-154-9</p> <p>Sequence 9, Application US/09911154</p> <p>Patent No. 6478809</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Harberd, Nicholas P</p> <p>APPLICANT: Peng, Jinrong</p> <p>APPLICANT: Carol, Pierre</p> <p>APPLICANT: Richards, Donald E</p> <p>TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana</p> <p>FILE REFERENCE: 620-158</p> <p>CURRENT APPLICATION NUMBER: US/09</p>									

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 1602 GGTATCGGTGGAGGAG 1619

RESULT 70

US-09-117-853-3
; Sequence 3, Application US/09117853
; Patent No. 6307126
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; CURRENT FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-117-853-3

Alignment Scores:
Pred. No.: 670 Length: 1643
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-117-853-3 (1-1643)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 1603 GGTATCGGTGGAGGAG 1620

RESULT 71

US-09-911-154-3
; Sequence 3, Application US/09911154
; Patent No. 6478809
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,154
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-154-3

Alignment Scores:
Pred. No.: 670 Length: 1643
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-911-154-3 (1-1643)

Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-911-154-3 (1-1643)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 1603 GGTATCGGTGGAGGAG 1620

RESULT 72

US-09-911-514-3
; Sequence 3, Application US/09911514
; Patent No. 6794560
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,514
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-911-514-3

Alignment Scores:

Pred. No.: 670 Length: 1643
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-911-514-3 (1-1643)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 1603 GGTATCGGTGGAGGAG 1620

RESULT 73

US-09-485-529-13
; Sequence 13, Application US/09485529
; Patent No. 6762348
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Richards, Donald E
; APPLICANT: Peng, Jinrong
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development
; FILE REFERENCE: 620-91
; CURRENT APPLICATION NUMBER: US/09/485,529
; CURRENT FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/GB98/02383
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: GB 9717192.0
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1768
; TYPE: DNA

```
; ORGANISM: Triticum aestivum
US-09-485-529-13
Alignment Scores:
Pred. No.: 728 Length: 1768
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-485-529-13 (1-1768)
Qy 1 GlyTyr***ValGluGlu 6
Db 1183 GGCTACCAAGTGGAGGAG 1200

RESULT 74
US-09-976-594-28/c
; Sequence 28, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 28
; LENGTH: 1771
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 230895.1
US-09-976-594-28
Alignment Scores:
Pred. No.: 729 Length: 1771
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-976-594-28 (1-1771)
Qy 1 GlyTyr***ValGluGlu 6
Db 873 GGATATGATGTAGAGAA 856

RESULT 75
US-08-750-134A-4/c
; Sequence 4, Application US/08750134A
; Patent No. 5985603
; GENERAL INFORMATION:
; APPLICANT: VALEIRA, SOLEDAD
; APPLICANT: BUELL, GARY
; TITLE OF INVENTION: P2X RECEPTORS (PURINOCEPTOR FAMILY)
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,134A
; FILING DATE: 22-JAN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR C.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 1430-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4006
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1837 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: rat P2x from vas deferens
US-08-750-134A-4
Alignment Scores:
Pred. No.: 760 Length: 1837
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 92.86% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-5 (1-6) x US-08-750-134A-4 (1-1837)
Qy 1 GlyTyr***ValGluGlu 6
Db 1750 GGCTACCATGTGGAAGAG 1733

Search completed: November 3, 2004, 17:56:24
Job time : 111.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2004, 16:31:21 ; Search time 274 Seconds
(without alignments)
112.283 Million cell updates/sec

Title: US-10-030-194A-5
Perfect score: 28
Sequence: 1 GYXVEE 6

Scoring table:
BLOSUMP62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10030194/runat_01112004.184849_29071/app_query.fasta_1.398
-DB=Published_Applications_NA -QfM=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-TRANS=human40.cdi -LIST=1000 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10030194 @CGN 1 1 1068 @runat_01112004.184849_29071
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	27	96.4	200	16	US-10-242-535A-58809	Sequence 58809, A
2	27	96.4	200	16	US-10-085-783A-58809	Sequence 58809, A
3	27	96.4	318	11	US-09-864-408A-3483	Sequence 3483, Ap
C 4	27	96.4	384	9	US-09-560-863-489	Sequence 269, App
C 5	27	96.4	394	17	US-10-767-701-30758	Sequence 30758, A
C 6	27	96.4	409	17	US-10-437-963-82192	Sequence 82192, A
C 7	27	96.4	417	10	US-09-918-995-4968	Sequence 4968, Ap
C 8	27	96.4	452	9	US-09-783-590-8368	Sequence 8368, Ap
C 9	27	96.4	475	17	US-10-021-323-11049	Sequence 11049, A
C 10	27	96.4	476	17	US-10-021-323-10992	Sequence 10992, A
C 11	27	96.4	479	9	US-09-974-300-4300	Sequence 4300, A
C 12	27	96.4	491	16	US-10-425-114-35722	Sequence 35722, A
C 13	27	96.4	498	9	US-09-917-800A-428	Sequence 428, App
C 14	27	96.4	535	9	US-09-974-300-1215	Sequence 1215, Ap
C 15	27	96.4	540	16	US-10-425-114-10779	Sequence 10779, A
C 16	27	96.4	543	17	US-10-021-323-10850	Sequence 10850, A
C 17	27	96.4	544	17	US-10-021-323-3889	Sequence 3889, Ap
C 18	27	96.4	545	17	US-10-021-323-11186	Sequence 11186, A
C 19	27	96.4	572	13	US-10-027-632-210714	Sequence 210714, A
C 20	27	96.4	572	15	US-10-027-632-210714	Sequence 210714, A
C 21	27	96.4	628	13	US-10-027-632-194951	Sequence 194951, A
C 22	27	96.4	628	15	US-10-027-632-194951	Sequence 194951, A
C 23	27	96.4	649	17	US-10-767-701-4268	Sequence 4268, Ap
C 24	27	96.4	659	13	US-10-027-632-8887	Sequence 8887, Ap
C 25	27	96.4	659	13	US-10-027-632-8888	Sequence 8888, Ap
C 26	27	96.4	659	13	US-10-027-632-8889	Sequence 8889, Ap
C 27	27	96.4	659	15	US-10-027-632-8888	Sequence 8887, Ap
C 28	27	96.4	659	15	US-10-027-632-8888	Sequence 8888, Ap
C 29	27	96.4	659	15	US-10-027-632-8888	Sequence 8889, Ap
C 30	27	96.4	664	13	US-10-027-632-226071	Sequence 226071, A
C 31	27	96.4	664	15	US-10-027-632-226071	Sequence 226071, A
C 32	27	96.4	747	9	US-09-764-887-485	Sequence 485, App
C 33	27	96.4	747	14	US-10-073-561-485	Sequence 485, App
C 34	27	96.4	759	16	US-10-424-599-51314	Sequence 51314, A
C 35	27	96.4	834	16	US-10-282-122A-6833	Sequence 6833, Ap
C 36	27	96.4	837	15	US-10-369-493-33361	Sequence 33361, A
C 37	27	96.4	942	13	US-10-369-493-23809	Sequence 23809, A
C 38	27	96.4	947	15	US-10-027-632-171842	Sequence 171842, A
C 39	27	96.4	947	15	US-10-027-632-171842	Sequence 171842, A
C 40	27	96.4	969	15	US-10-369-493-45793	Sequence 45793, A
C 41	27	96.4	978	17	US-10-767-795-1815	Sequence 1815, Ap
C 42	27	96.4	987	9	US-09-886-055-302	Sequence 302, App
C 43	27	96.4	987	10	US-09-804-291-302	Sequence 302, App
C 44	27	96.4	987	15	US-10-387-629-109	Sequence 109, App
C 45	27	96.4	987	16	US-10-237-021-44	Sequence 44, Appl
C 46	27	96.4	1002	16	US-10-343-650A-361	Sequence 361, App
C 47	27	96.4	1017	17	US-10-024-399-31	Sequence 31, Appl
C 48	27	96.4	1017	17	US-10-437-963-73808	Sequence 73808, A
C 49	27	96.4	1035	9	US-09-815-242-4004	Sequence 4004, Ap
C 50	27	96.4	1035	16	US-10-282-122A-7327	Sequence 7327, Ap
C 51	27	96.4	1099	15	US-10-369-493-40154	Sequence 40154, A
C 52	27	96.4	1110	17	US-10-767-795-1814	Sequence 1814, Ap
C 53	27	96.4	1269	16	US-10-260-238-779	Sequence 779, App
C 54	27	96.4	1269	17	US-10-437-963-15300	Sequence 15300, A
C 55	27	96.4	1282	16	US-10-425-114-20541	Sequence 20541, A
C 56	27	96.4	1362	17	US-10-437-963-25616	Sequence 25616, A
C 57	27	96.4	1378	15	US-10-369-493-36451	Sequence 36451, A
C 58	27	96.4	1387	15	US-10-017-161-285	Sequence 285, App
C 59	27	96.4	1387	15	US-10-292-798-255	Sequence 255, App
C 60	27	96.4	1425	16	US-10-354-437-31	Sequence 31, Appl
C 61	27	96.4	1635	9	US-09-815-242-4536	Sequence 4536, Ap
C 62	27	96.4	1662	9	US-09-815-242-8100	Sequence 8100, Ap
C 63	27	96.4	1662	16	US-10-282-122A-7877	Sequence 7877, Ap
C 64	27	96.4	1728	17	US-10-437-963-70740	Sequence 70740, A
C 65	27	96.4	1781	16	US-10-424-559-24574	Sequence 24574, A
C 66	27	96.4	1796	16	US-10-424-559-59017	Sequence 59017, A
C 67	27	96.4	1949	15	US-10-425-114-29590	Sequence 29590, A
C 68	27	96.4	1976	13	US-10-027-632-99250	Sequence 99250, A
C 69	27	96.4	1976	15	US-10-027-632-99250	Sequence 99250, A
C 70	27	96.4	2019	16	US-10-424-599-77544	Sequence 77544, A

c 71	27	96.4	2104	16	US-10-108-260A-2081	Sequence 2081, Ap	144	26	92.9	442	16	US-10-424-599-55327	Sequence 55327, A
c 72	27	96.4	2151	15	US-10-369-493-33444	Sequence 33444, A	145	26	92.9	444	9	US-09-864-761-24780	Sequence 24780, A
c 73	27	96.4	2226	15	US-10-369-493-42642	Sequence 42642, A	146	26	92.9	445	13	US-10-027-450-12	Sequence 12, Appl
c 74	27	96.4	2289	9	US-09-391-340-9	Sequence 9, Appl	147	26	92.9	455	13	US-10-027-632-88245	Sequence 88245, A
c 75	27	96.4	2289	9	US-09-948-369-9	Sequence 9, Appl	148	26	92.9	455	15	US-10-027-632-88245	Sequence 88245, A
c 76	27	96.4	2309	16	US-10-424-599-109493	Sequence 109493, A	149	26	92.9	456	9	US-09-864-761-6192	Sequence 6192, Ap
c 77	27	96.4	2682	9	US-09-822-830A-228	Sequence 228, App	150	26	92.9	461	10	US-09-918-995-790	Sequence 790, App
c 78	27	96.4	3040	16	US-10-062-674-2172	Sequence 2172, Ap	151	26	92.9	461	17	US-10-021-323-5756	Sequence 5756, App
c 79	27	96.4	3962	18	US-10-416-330-2	Sequence 2, Appl	152	26	92.9	464	9	US-09-925-300-659	Sequence 659, App
c 80	27	96.4	4217	16	US-10-282-122A-9845	Sequence 9845, Ap	153	26	92.9	466	9	US-09-864-761-385	Sequence 385, App
c 81	27	96.4	5925	15	US-10-311-455-1549	Sequence 1549, Ap	154	26	92.9	466	10	US-09-918-995-13435	Sequence 13435, A
c 82	27	96.4	6571	9	US-09-969-347-172	Sequence 172, App	155	26	92.9	474	13	US-10-027-632-56422	Sequence 56422, A
c 83	27	96.4	6571	9	US-09-971-392-95	Sequence 95, Appl	156	26	92.9	474	13	US-10-027-632-179659	Sequence 179659, A
c 84	27	96.4	13321	8	US-08-781-986A-4	Sequence 4, Appl	157	26	92.9	474	13	US-10-027-632-310751	Sequence 310751, A
c 85	27	96.4	13321	16	US-10-329-624-4	Sequence 4, Appl	158	26	92.9	474	15	US-10-027-632-56422	Sequence 56422, A
c 86	27	96.4	17276	9	US-09-870-759-83	Sequence 83, Appl	159	26	92.9	474	15	US-10-027-632-179659	Sequence 179659, A
c 87	27	96.4	17276	10	US-09-751-708A-83	Sequence 83, Appl	160	26	92.9	474	15	US-10-027-632-310751	Sequence 310751, A
c 88	27	96.4	25020	16	US-10-192-280-1	Sequence 1, Appl	161	26	92.9	474	16	US-10-242-535A-28625	Sequence 28625, A
c 89	27	96.4	49979	13	US-10-741-601-5746	Sequence 5746, Ap	162	26	92.9	474	16	US-10-085-783A-28625	Sequence 28625, A
c 90	27	96.4	83888	13	US-10-087-192-541	Sequence 541, App	163	26	92.9	477	13	US-10-027-632-57169	Sequence 57169, A
c 91	27	96.4	96598	16	US-10-394-948-25	Sequence 25, Appl	164	26	92.9	477	15	US-10-027-632-57169	Sequence 57169, A
c 92	27	96.4	96598	16	US-10-052-482-115	Sequence 115, App	165	26	92.9	480	9	US-09-938-842A-2613	Sequence 2613, Ap
c 93	27	96.4	101046	17	US-10-741-601-5689	Sequence 5689, Ap	166	26	92.9	480	11	US-09-338-842A-2613	Sequence 2613, Ap
c 94	27	96.4	370469	13	US-10-087-192-250	Sequence 250, App	167	26	92.9	484	16	US-10-424-599-73523	Sequence 73523, A
c 95	26	92.9	65	10	US-09-908-975-28104	Sequence 28104, A	168	26	92.9	486	16	US-10-398-221-770	Sequence 770, App
c 96	26	92.9	150	9	US-09-864-761-17186	Sequence 17186, A	169	26	92.9	486	16	US-10-424-599-48422	Sequence 48422, A
c 97	26	92.9	165	9	US-09-864-761-22931	Sequence 22931, A	170	26	92.9	496	16	US-10-424-599-48422	Sequence 48422, A
c 98	26	92.9	165	15	US-10-029-386-14073	Sequence 14073, A	171	26	92.9	499	10	US-09-814-353-6167	Sequence 6167, Ap
c 99	26	92.9	179	15	US-10-029-386-19359	Sequence 19359, A	172	26	92.9	499	10	US-09-814-353-12445	Sequence 12445, A
c 100	26	92.9	201	17	US-10-741-601-8712	Sequence 8712, Ap	173	26	92.9	500	11	US-09-732-627A-2018	Sequence 2018, Ap
c 101	26	92.9	201	17	US-10-741-601-9848	Sequence 9848, Ap	174	26	92.9	511	16	US-10-424-599-75777	Sequence 75777, A
c 102	26	92.9	201	17	US-10-741-601-12768	Sequence 12768, Ap	175	26	92.9	511	16	US-10-424-599-107511	Sequence 107511, A
c 103	26	92.9	201	17	US-10-741-601-13419	Sequence 13419, A	176	26	92.9	515	15	US-10-029-386-5915	Sequence 5915, Ap
c 104	26	92.9	206	17	US-10-437-963-90334	Sequence 90334, A	177	26	92.9	519	9	US-09-864-761-8349	Sequence 8349, Ap
c 105	26	92.9	217	16	US-10-424-599-94659	Sequence 94659, A	178	26	92.9	519	15	US-10-029-386-24183	Sequence 24183, A
c 106	26	92.9	234	9	US-09-728-445-614	Sequence 614, App	179	26	92.9	523	9	US-09-864-761-8787	Sequence 8787, Ap
c 107	26	92.9	239	16	US-10-242-535A-27907	Sequence 27907, A	180	26	92.9	537	15	US-10-029-386-1161	Sequence 1161, Ap
c 108	26	92.9	239	16	US-10-085-783A-27907	Sequence 27907, A	181	26	92.9	539	15	US-10-029-386-8458	Sequence 8458, Ap
c 109	26	92.9	246	17	US-10-437-963-22583	Sequence 22583, A	182	26	92.9	542	9	US-09-864-761-8938	Sequence 8938, Ap
c 110	26	92.9	273	9	US-09-864-761-25080	Sequence 25080, A	183	26	92.9	543	9	US-09-864-761-8034	Sequence 8034, Ap
c 111	26	92.9	274	15	US-10-029-386-22158	Sequence 22158, A	184	26	92.9	547	17	US-10-437-963-93308	Sequence 93308, A
c 112	26	92.9	283	15	US-10-029-386-19646	Sequence 19646, A	185	26	92.9	547	15	US-10-029-386-10482	Sequence 10482, A
c 113	26	92.9	287	15	US-10-029-386-13874	Sequence 13874, A	186	26	92.9	547	16	US-10-424-599-42086	Sequence 42086, A
c 114	26	92.9	288	16	US-10-335-977-3537	Sequence 3537, Ap	187	26	92.9	549	9	US-09-864-761-8284	Sequence 8284, A
c 115	26	92.9	296	16	US-10-424-599-34787	Sequence 34787, A	188	26	92.9	550	9	US-09-770-152-540	Sequence 540, App
c 116	26	92.9	305	10	US-09-933-261-4	Sequence 4, Appl	189	26	92.9	552	17	US-10-021-323-4051	Sequence 4051, Ap
c 117	26	92.9	305	14	US-10-256-702-4	Sequence 4, Appl	190	26	92.9	553	13	US-10-027-632-211235	Sequence 211235, A
c 118	26	92.9	336	11	US-09-864-408A-3457	Sequence 3457, Ap	191	26	92.9	553	13	US-10-027-632-211236	Sequence 211236, A
c 119	26	92.9	343	9	US-09-864-761-25484	Sequence 25484, A	192	26	92.9	553	15	US-10-027-632-211235	Sequence 211235, A
c 120	26	92.9	347	16	US-10-242-535A-27571	Sequence 27571, A	193	26	92.9	553	15	US-10-027-632-211236	Sequence 211236, A
c 121	26	92.9	347	16	US-10-085-783A-27571	Sequence 27571, A	194	26	92.9	554	9	US-09-924-035A-622	Sequence 622, App
c 122	26	92.9	355	16	US-10-424-599-12515	Sequence 12515, A	195	26	92.9	558	15	US-10-029-386-12199	Sequence 12199, A
c 123	26	92.9	357	16	US-10-335-977-3538	Sequence 3538, Ap	196	26	92.9	559	10	US-09-918-995-28035	Sequence 28035, A
c 124	26	92.9	369	17	US-10-437-963-31269	Sequence 31269, A	197	26	92.9	559	13	US-10-027-632-205208	Sequence 205208, A
c 125	26	92.9	370	16	US-10-424-599-129175	Sequence 129175, A	198	26	92.9	559	13	US-10-027-632-205209	Sequence 205209, A
c 126	26	92.9	382	9	US-09-770-791-152	Sequence 152, App	199	26	92.9	559	15	US-10-027-632-205208	Sequence 205208, A
c 127	26	92.9	388	9	US-09-783-590-8848	Sequence 8848, Ap	200	26	92.9	559	15	US-10-027-632-205209	Sequence 205209, A
c 128	26	92.9	394	9	US-09-864-761-18352	Sequence 18352, A	201	26	92.9	559	15	US-10-027-632-320581	Sequence 320581, A
c 129	26	92.9	394	16	US-10-424-599-17096	Sequence 17096, A	202	26	92.9	559	15	US-10-029-386-4517	Sequence 4517, Ap
c 130	26	92.9	406	17	US-10-430-201-2200	Sequence 2200, Ap	203	26	92.9	563	15	US-10-029-386-167	Sequence 167, App
c 131	26	92.9	406	17	US-10-430-201-2201	Sequence 2201, Ap	204	26	92.9	566	15	US-10-029-386-167	Sequence 167, App
c 132	26	92.9	407	15	US-10-029-386-25901	Sequence 25901, A	205	26	92.9	566	17	US-10-021-323-12171	Sequence 12171, A
c 133	26	92.9	408	9	US-09-864-761-25017	Sequence 25017, A	206	26	92.9	567	10	US-09-814-353-21918	Sequence 21918, A
c 134	26	92.9	409	17	US-10-437-963-64680	Sequence 64680, A	207	26	92.9	569	16	US-10-425-114-31715	Sequence 31715, A
c 135	26	92.9	430	15	US-10-029-386-18372	Sequence 18372, A	208	26	92.9	570	9	US-09-864-761-8811	Sequence 8811, Ap
c 136	26	92.9	431	11	US-09-987-899-4903	Sequence 4903, Ap	209	26	92.9	570	9	US-09-864-761-8817	Sequence 8817, Ap
c 137	26	92.9	435	9	US-09-764-877-311	Sequence 311, App	210	26	92.9	571	15	US-10-029-386-368	Sequence 368, App
c 138	26	92.9	435	16	US-10-242-515-311	Sequence 311, App	211	26	92.9	573	9	US-09-974-300-5346	Sequence 5346, Ap
c 139	26	92.9	437	13	US-10-027-632-2817	Sequence 2817, Ap	212	26	92.9	575	10	US-09-933-767-91	Sequence 91, Appl
c 140	26	92.9	437	15	US-10-029-386-18273	Sequence 18273, A	213	26	92.9	575	14	US-10-004-860-91	Sequence 91, Appl
c 141	26	92.9	437	15	US-10-027-632-2817	Sequence 2817, Ap	214	26	92.9	575	14	US-10-023-282-91	Sequence 91, Appl
c 142	26	92.9	440	17	US-10-767-795-3853	Sequence 3853, A	215	26	92.9	576	9	US-09-770-152-212	Sequence 212, App
c 143	26	92.9	442	16	US-10-424-599-27109	Sequence 27109, A	216	26	92.9	576	9	US-09-864-761-9294	Sequence 9294, Ap

c 217	26	92.9	580	9	US-09-864-761-9802	Sequence 9802, Ap	c 230	26	92.9	711	16	US-10-267-502-63	Sequence 63, Appl
218	26	92.9	580	13	US-10-027-632-321582	Sequence 321582, Ap	231	26	92.9	712	16	US-10-424-599-127380	Sequence 127380, Ap
219	26	92.9	580	13	US-10-027-632-321583	Sequence 321583, Ap	232	26	92.9	713	16	US-10-027-632-3203	Sequence 3203, Ap
220	26	92.9	580	15	US-10-027-632-321582	Sequence 321582, Ap	233	26	92.9	713	15	US-10-027-632-3203	Sequence 3203, Ap
221	26	92.9	580	15	US-10-027-632-321583	Sequence 321583, Ap	234	26	92.9	721	16	US-10-276-774-188	Sequence 188, App
222	26	92.9	583	15	US-10-410-681-19	Sequence 19, Appl	235	26	92.9	721	16	US-10-027-632-153692	Sequence 153692, Ap
223	26	92.9	584	9	US-09-864-761-8921	Sequence 8921, Ap	236	26	92.9	752	13	US-10-027-632-153692	Sequence 153692, Ap
224	26	92.9	585	13	US-10-027-632-205297	Sequence 205297, Ap	237	26	92.9	752	13	US-10-027-632-153694	Sequence 153694, Ap
225	26	92.9	585	13	US-10-027-632-229284	Sequence 229284, Ap	238	26	92.9	752	13	US-10-027-632-153692	Sequence 153692, Ap
226	26	92.9	585	13	US-10-027-632-229285	Sequence 229285, Ap	239	26	92.9	752	15	US-10-027-632-153692	Sequence 153692, Ap
227	26	92.9	585	15	US-10-027-632-205297	Sequence 205297, Ap	240	26	92.9	752	15	US-10-027-632-153694	Sequence 153694, Ap
228	26	92.9	585	15	US-10-027-632-229284	Sequence 229284, Ap	241	26	92.9	782	15	US-10-027-632-153694	Sequence 21, Appl
229	26	92.9	585	15	US-10-027-632-229285	Sequence 229285, Ap	242	26	92.9	782	15	US-10-410-681-21	Sequence 153653, Ap
230	26	92.9	589	16	US-10-027-632-229285	Sequence 229285, Ap	243	26	92.9	786	13	US-10-027-632-153653	Sequence 153653, Ap
231	26	92.9	589	16	US-10-029-386-4616	Sequence 4616, Ap	244	26	92.9	791	13	US-10-027-632-153653	Sequence 153653, Ap
232	26	92.9	590	15	US-10-029-386-5603	Sequence 5603, Ap	245	26	92.9	791	13	US-10-027-632-156321	Sequence 156321, Ap
233	26	92.9	591	13	US-10-027-632-221393	Sequence 221393, Ap	246	26	92.9	791	15	US-10-027-632-156321	Sequence 156321, Ap
234	26	92.9	591	15	US-10-027-632-221393	Sequence 221393, Ap	247	26	92.9	810	16	US-10-424-599-24572	Sequence 24572, A
235	26	92.9	598	13	US-10-027-632-186972	Sequence 186972, Ap	248	26	92.9	837	17	US-10-767-701-11942	Sequence 11942, A
236	26	92.9	598	13	US-10-027-632-186972	Sequence 186972, Ap	249	26	92.9	837	15	US-10-369-493-27755	Sequence 27755, A
237	26	92.9	604	9	US-09-864-761-17057	Sequence 17057, Ap	250	26	92.9	855	13	US-10-027-632-170302	Sequence 170302, Ap
238	26	92.9	604	9	US-09-864-761-17057	Sequence 17057, Ap	251	26	92.9	855	13	US-10-027-632-170303	Sequence 170303, Ap
239	26	92.9	608	13	US-10-027-632-308484	Sequence 308484, Ap	252	26	92.9	855	13	US-10-027-632-170302	Sequence 170302, Ap
240	26	92.9	608	13	US-10-027-632-308484	Sequence 308484, Ap	253	26	92.9	855	15	US-10-027-632-170303	Sequence 170303, Ap
241	26	92.9	611	13	US-10-027-632-213303	Sequence 213303, Ap	254	26	92.9	855	15	US-10-027-632-170304	Sequence 170304, Ap
242	26	92.9	611	13	US-10-027-632-222790	Sequence 222790, Ap	255	26	92.9	861	15	US-10-029-386-24197	Sequence 24197, A
243	26	92.9	611	13	US-10-027-632-241898	Sequence 241898, Ap	256	26	92.9	876	16	US-10-424-599-6121	Sequence 6121, Ap
244	26	92.9	611	15	US-10-027-632-213303	Sequence 213303, Ap	257	26	92.9	909	15	US-10-156-761-4077	Sequence 4077, Ap
245	26	92.9	611	15	US-10-027-632-222790	Sequence 222790, Ap	258	26	92.9	916	16	US-10-424-599-60410	Sequence 60410, A
246	26	92.9	611	15	US-10-027-632-241898	Sequence 241898, Ap	259	26	92.9	932	13	US-10-027-632-136343	Sequence 136343, Ap
247	26	92.9	612	13	US-10-027-632-242079	Sequence 242079, Ap	260	26	92.9	932	15	US-10-027-632-136343	Sequence 136343, Ap
248	26	92.9	612	13	US-10-027-632-242079	Sequence 242079, Ap	261	26	92.9	935	13	US-10-027-632-163064	Sequence 163064, Ap
249	26	92.9	615	13	US-10-027-632-4545	Sequence 4545, Ap	262	26	92.9	935	15	US-10-027-632-163064	Sequence 163064, Ap
250	26	92.9	615	15	US-10-027-632-4545	Sequence 4545, Ap	263	26	92.9	981	16	US-10-282-122A-15227	Sequence 15227, A
251	26	92.9	619	17	US-10-767-701-9112	Sequence 9112, Ap	264	26	92.9	994	17	US-10-425-114-8439	Sequence 8439, Ap
252	26	92.9	620	13	US-10-027-632-239079	Sequence 239079, Ap	265	26	92.9	994	17	US-10-767-701-12175	Sequence 12175, A
253	26	92.9	620	13	US-10-027-632-239080	Sequence 239080, Ap	266	26	92.9	1000	13	US-10-027-632-250374	Sequence 250374, Ap
254	26	92.9	620	15	US-10-027-632-239079	Sequence 239079, Ap	267	26	92.9	1000	15	US-10-027-632-250374	Sequence 250374, Ap
255	26	92.9	620	15	US-10-027-632-239080	Sequence 239080, Ap	268	26	92.9	1001	15	US-10-170-097-204	Sequence 204, App
256	26	92.9	622	16	US-10-424-599-125715	Sequence 125715, Ap	269	26	92.9	1009	16	US-10-425-114-29334	Sequence 29334, A
257	26	92.9	622	16	US-10-424-599-125715	Sequence 125715, Ap	270	26	92.9	1009	16	US-10-425-114-29347	Sequence 29347, A
258	26	92.9	623	15	US-10-027-632-268485	Sequence 268485, Ap	271	26	92.9	1009	16	US-10-027-632-293342	Sequence 293342, Ap
259	26	92.9	624	13	US-10-027-632-201623	Sequence 201623, Ap	272	26	92.9	1025	15	US-10-027-632-293342	Sequence 293342, Ap
260	26	92.9	624	13	US-10-027-632-201623	Sequence 201623, Ap	273	26	92.9	1025	15	US-09-764-870-171	Sequence 171, App
261	26	92.9	625	13	US-10-401-343-9	Sequence 9, Appl	274	26	92.9	1026	14	US-10-125-540-171	Sequence 171, App
262	26	92.9	635	13	US-10-027-632-267941	Sequence 267941, Ap	275	26	92.9	1026	14	US-10-369-493-40329	Sequence 40329, A
263	26	92.9	635	15	US-10-027-632-267941	Sequence 267941, Ap	276	26	92.9	1035	15	US-10-369-493-32555	Sequence 32555, A
264	26	92.9	640	13	US-10-027-632-230373	Sequence 230373, Ap	277	26	92.9	1044	15	US-10-369-493-32555	Sequence 15128, A
265	26	92.9	640	13	US-10-027-632-230373	Sequence 230373, Ap	278	26	92.9	1044	16	US-10-425-114-15128	Sequence 18690, A
266	26	92.9	655	17	US-10-021-323-5665	Sequence 5665, Ap	279	26	92.9	1047	16	US-10-425-114-18690	Sequence 16027, A
267	26	92.9	658	17	US-10-767-701-8545	Sequence 8545, Ap	280	26	92.9	1048	16	US-10-424-599-16027	Sequence 50, Appl
268	26	92.9	663	13	US-10-027-632-211764	Sequence 211764, Ap	281	26	92.9	1056	18	US-10-633-850-51	Sequence 51, Appl
269	26	92.9	663	13	US-10-027-632-211765	Sequence 211765, Ap	282	26	92.9	1056	18	US-10-633-850-51	Sequence 53, Appl
270	26	92.9	663	13	US-10-027-632-211766	Sequence 211766, Ap	283	26	92.9	1056	18	US-10-633-850-55	Sequence 55, Appl
271	26	92.9	663	13	US-10-027-632-211767	Sequence 211767, Ap	284	26	92.9	1056	18	US-10-633-850-57	Sequence 57, Appl
272	26	92.9	663	13	US-10-027-632-211767	Sequence 211767, Ap	285	26	92.9	1056	18	US-10-633-850-59	Sequence 59, Appl
273	26	92.9	663	15	US-10-027-632-211765	Sequence 211765, Ap	286	26	92.9	1056	18	US-10-633-850-61	Sequence 61, Appl
274	26	92.9	663	15	US-10-027-632-211766	Sequence 211766, Ap	287	26	92.9	1056	18	US-10-633-850-63	Sequence 63, Appl
275	26	92.9	663	15	US-10-027-632-211767	Sequence 211767, Ap	288	26	92.9	1056	18	US-10-633-850-65	Sequence 65, Appl
276	26	92.9	668	10	US-09-814-353-18829	Sequence 18829, A	289	26	92.9	1056	18	US-10-633-850-67	Sequence 67, Appl
277	26	92.9	670	13	US-10-027-632-225247	Sequence 225247, Ap	290	26	92.9	1056	18	US-10-633-850-69	Sequence 69, Appl
278	26	92.9	670	13	US-10-027-632-225248	Sequence 225248, Ap	291	26	92.9	1056	18	US-10-633-850-71	Sequence 71, Appl
279	26	92.9	670	15	US-10-027-632-225247	Sequence 225247, Ap	292	26	92.9	1056	18	US-10-633-850-73	Sequence 73, Appl
280	26	92.9	670	15	US-10-027-632-225248	Sequence 225248, Ap	293	26	92.9	1056	18	US-10-633-850-75	Sequence 75, Appl
281	26	92.9	678	16	US-10-424-599-89179	Sequence 89179, A	294	26	92.9	1056	18	US-10-633-850-77	Sequence 77, Appl
282	26	92.9	685	13	US-10-027-632-19666	Sequence 19666, A	295	26	92.9	1056	18	US-10-633-850-79	Sequence 79, Appl
283	26	92.9	685	13	US-10-027-632-19666	Sequence 19666, A	296	26	92.9	1056	18	US-10-633-850-81	Sequence 81, Appl
284	26	92.9	695	13	US-10-027-632-121602	Sequence 121602, Ap	297	26	92.9	1056	18	US-10-633-850-83	Sequence 83, Appl
285	26	92.9	695	13	US-10-027-632-121603	Sequence 121603, Ap	298	26	92.9	1056	18	US-10-633-850-85	Sequence 85, Appl
286	26	92.9	695	15	US-10-027-632-121602	Sequence 121602, Ap	299	26	92.9	1056	18	US-10-633-850-87	Sequence 87, Appl
287	26	92.9	706	13	US-10-027-632-121603	Sequence 121603, Ap	300	26	92.9	1056	18	US-10-633-850-89	Sequence 89, Appl
288	26	92.9	706	13	US-10-027-632-110864	Sequence 110864, Ap	301	26	92.9	1056	18	US-10-633-850-91	Sequence 91, Appl
289	26	92.9	706	15	US-10-027-632-110864	Sequence 110864, Ap	302	26	92.9	1056	18	US-10-633-850-93	Sequence 93, Appl

C 363	26	92.9	1056	18	US-10-633-850-95	Sequence 95, Appl	436	26	92.9	1636	9	US-09-911-513-7	Sequence 7, Appl
C 364	26	92.9	1056	18	US-10-633-850-97	Sequence 97, Appl	437	26	92.9	1636	10	US-09-911-514-7	Sequence 7, Appl
C 365	26	92.9	1059	13	US-10-027-632-265494	Sequence 265494, A	438	26	92.9	1642	9	US-09-911-513-5	Sequence 5, Appl
C 366	26	92.9	1059	13	US-10-027-632-265494	Sequence 265494, A	439	26	92.9	1642	9	US-09-911-513-9	Sequence 9, Appl
C 367	26	92.9	1063	16	US-10-424-599-7718	Sequence 7718, A	440	26	92.9	1642	10	US-09-911-514-5	Sequence 5, Appl
C 368	26	92.9	1070	16	US-10-425-114-19077	Sequence 19077, A	441	26	92.9	1642	10	US-09-911-514-9	Sequence 9, Appl
C 369	26	92.9	1074	16	US-10-282-122A-34287	Sequence 34287, A	442	26	92.9	1643	9	US-09-911-513-3	Sequence 3, Appl
C 370	26	92.9	1074	16	US-10-282-122A-35392	Sequence 35392, A	443	26	92.9	1643	10	US-09-911-514-3	Sequence 3, Appl
C 371	26	92.9	1075	16	US-10-425-114-20656	Sequence 20656, A	C 444	26	92.9	1667	16	US-10-398-221-3441	Sequence 3441, A
C 372	26	92.9	1077	16	US-10-282-122A-34066	Sequence 34066, A	C 445	26	92.9	1671	15	US-10-292-896-61	Sequence 61, A
C 373	26	92.9	1080	16	US-10-425-114-14907	Sequence 14907, A	446	26	92.9	1678	16	US-10-296-115-65	Sequence 65, A
C 374	26	92.9	1086	13	US-10-027-450-14	Sequence 14, Appl	C 447	26	92.9	1690	17	US-10-437-963-23377	Sequence 23377, A
C 375	26	92.9	1086	13	US-10-369-493-24316	Sequence 24316, A	C 448	26	92.9	1695	16	US-10-425-114-11671	Sequence 11671, A
C 376	26	92.9	1086	13	US-10-369-493-26767	Sequence 26767, A	C 449	26	92.9	1707	16	US-10-424-599-128857	Sequence 128857, A
C 377	26	92.9	1097	13	US-10-027-632-31507	Sequence 31507, A	450	26	92.9	1710	16	US-10-425-114-11778	Sequence 11778, A
C 378	26	92.9	1097	13	US-10-027-632-31507	Sequence 31507, A	451	26	92.9	1728	16	US-10-425-114-8187	Sequence 8187, A
C 379	26	92.9	1102	16	US-10-424-599-9756	Sequence 9756, A	452	26	92.9	1734	9	US-09-805-204-3	Sequence 3, Appl
C 380	26	92.9	1109	16	US-10-398-221-1306	Sequence 1306, A	453	26	92.9	1734	13	US-10-103-511-3	Sequence 3, Appl
C 381	26	92.9	1110	15	US-10-369-493-34995	Sequence 34995, A	454	26	92.9	1764	9	US-09-938-842A-2513	Sequence 2513, A
C 382	26	92.9	1137	9	US-09-738-626-2414	Sequence 2414, A	455	26	92.9	1764	11	US-09-938-842A-2513	Sequence 2513, A
C 383	26	92.9	1145	8	US-08-838-151A-19	Sequence 19, Appl	456	26	92.9	1773	16	US-10-412-6998-215	Sequence 215, A
C 384	26	92.9	1145	8	US-08-838-151A-23	Sequence 23, Appl	C 457	26	92.9	1773	9	US-09-938-842A-1218	Sequence 1218, A
C 385	26	92.9	1145	8	US-08-838-151A-26	Sequence 26, Appl	C 458	26	92.9	1773	11	US-09-938-842A-1218	Sequence 1218, A
C 386	26	92.9	1145	8	US-08-838-151A-29	Sequence 29, Appl	459	26	92.9	1788	17	US-10-437-963-73600	Sequence 73600, A
C 387	26	92.9	1153	17	US-10-767-701-2260	Sequence 2260, A	C 460	26	92.9	1836	16	US-10-282-122A-12610	Sequence 12610, A
C 388	26	92.9	1155	9	US-09-938-842A-1230	Sequence 1230, A	461	26	92.9	1845	15	US-10-369-493-37936	Sequence 37936, A
C 389	26	92.9	1155	11	US-09-938-842A-1230	Sequence 1230, A	462	26	92.9	1889	16	US-10-425-114-12943	Sequence 12943, A
C 390	26	92.9	1170	16	US-10-425-114-34558	Sequence 34558, A	463	26	92.9	1917	16	US-10-467-433-39	Sequence 39, A
C 391	26	92.9	1173	17	US-10-437-963-10421	Sequence 10421, A	C 464	26	92.9	1927	15	US-10-225-068-71	Sequence 71, A
C 392	26	92.9	1178	13	US-10-027-632-206612	Sequence 206612, A	C 465	26	92.9	1927	15	US-10-225-068A-909	Sequence 909, A
C 393	26	92.9	1178	13	US-10-027-632-206612	Sequence 206612, A	C 466	26	92.9	1927	16	US-10-374-780A-353	Sequence 353, A
C 394	26	92.9	1190	17	US-10-027-632-206612	Sequence 206612, A	467	26	92.9	1929	16	US-10-398-221-1939	Sequence 1939, A
C 395	26	92.9	1197	13	US-10-027-632-216823	Sequence 216823, A	468	26	92.9	1945	9	US-09-864-761-2925	Sequence 2925, A
C 396	26	92.9	1197	13	US-10-027-632-216823	Sequence 216823, A	469	26	92.9	1951	15	US-10-278-536-45	Sequence 45, A
C 397	26	92.9	1230	15	US-10-369-493-43541	Sequence 43541, A	470	26	92.9	1951	15	US-10-225-066A-783	Sequence 783, A
C 398	26	92.9	1238	13	US-10-027-632-123862	Sequence 123862, A	471	26	92.9	1951	16	US-10-374-780A-2289	Sequence 2289, A
C 399	26	92.9	1238	13	US-10-027-632-123862	Sequence 123862, A	472	26	92.9	1951	16	US-10-412-6998-217	Sequence 217, A
C 400	26	92.9	1304	15	US-10-424-599-7720	Sequence 7720, A	473	26	92.9	1956	16	US-10-424-599-65202	Sequence 65202, A
C 401	26	92.9	1305	15	US-10-232-798-507	Sequence 507, A	474	26	92.9	1964	9	US-09-911-513-1	Sequence 1, A
C 402	26	92.9	1358	13	US-10-027-632-43913	Sequence 43913, A	475	26	92.9	1964	10	US-09-911-514-1	Sequence 1, A
C 403	26	92.9	1358	13	US-10-027-632-43913	Sequence 43913, A	476	26	92.9	1975	9	US-09-822-849A-202	Sequence 202, A
C 404	26	92.9	1362	16	US-10-425-114-17415	Sequence 17415, A	C 477	26	92.9	1975	13	US-10-027-632-97432	Sequence 97432, A
C 405	26	92.9	1383	16	US-10-425-114-11532	Sequence 11532, A	C 478	26	92.9	1976	13	US-10-027-632-97432	Sequence 97432, A
C 406	26	92.9	1397	16	US-10-424-599-99057	Sequence 99057, A	C 479	26	92.9	1976	15	US-10-027-632-97432	Sequence 97432, A
C 407	26	92.9	1416	13	US-10-027-632-49085	Sequence 49085, A	C 480	26	92.9	1976	15	US-10-027-632-97432	Sequence 97432, A
C 408	26	92.9	1416	13	US-10-027-632-63574	Sequence 63574, A	481	26	92.9	2008	16	US-10-398-221-1765	Sequence 1765, A
C 409	26	92.9	1416	13	US-10-027-632-49085	Sequence 49085, A	482	26	92.9	2008	16	US-10-398-221-3414	Sequence 3414, A
C 410	26	92.9	1416	16	US-10-027-632-63574	Sequence 63574, A	483	26	92.9	2028	17	US-10-767-795-3851	Sequence 3851, A
C 411	26	92.9	1418	16	US-10-425-114-14732	Sequence 14732, A	484	26	92.9	2064	17	US-10-437-963-67833	Sequence 67833, A
C 412	26	92.9	1452	15	US-10-369-493-41950	Sequence 41950, A	485	26	92.9	2091	15	US-10-104-047-1673	Sequence 1673, A
C 413	26	92.9	1453	15	US-10-020-513-1	Sequence 1, Appl	486	26	92.9	2106	13	US-10-002-600-6	Sequence 6, A
C 414	26	92.9	1484	16	US-10-425-114-25171	Sequence 25171, A	C 487	26	92.9	2125	17	US-10-437-963-92841	Sequence 92841, A
C 415	26	92.9	1506	16	US-10-467-433-21	Sequence 21, Appl	488	26	92.9	2133	17	US-10-437-963-68845	Sequence 68845, A
C 416	26	92.9	1506	17	US-10-450-681A-1	Sequence 1, Appl	489	26	92.9	2159	17	US-10-781-014-595	Sequence 595, A
C 417	26	92.9	1524	16	US-10-282-122A-13650	Sequence 13650, A	490	26	92.9	2174	16	US-10-424-599-8444	Sequence 8444, A
C 418	26	92.9	1539	17	US-10-437-963-69362	Sequence 69362, A	491	26	92.9	2184	9	US-09-938-842A-2661	Sequence 2661, A
C 419	26	92.9	1549	16	US-09-374-046A-41	Sequence 41, Appl	492	26	92.9	2184	11	US-09-938-842A-2661	Sequence 2661, A
C 420	26	92.9	1549	16	US-10-616-263-41	Sequence 41, Appl	C 493	26	92.9	2217	9	US-09-738-626-3488	Sequence 3488, A
C 421	26	92.9	1552	16	US-10-424-599-13193	Sequence 13193, A	C 494	26	92.9	2233	9	US-09-817-913-10	Sequence 10, A
C 422	26	92.9	1553	16	US-10-424-599-31264	Sequence 31264, A	C 495	26	92.9	2233	9	US-09-817-913-10	Sequence 10, A
C 423	26	92.9	1566	13	US-10-027-632-65358	Sequence 65358, A	C 496	26	92.9	2233	10	US-09-563-728A-33	Sequence 33, A
C 424	26	92.9	1566	15	US-10-027-632-65358	Sequence 65358, A	C 497	26	92.9	2240	16	US-10-351-334-84	Sequence 84, A
C 425	26	92.9	1568	16	US-10-425-114-31185	Sequence 31185, A	498	26	92.9	2264	17	US-10-465-498-90	Sequence 90, A
C 426	26	92.9	1602	9	US-09-938-842A-1026	Sequence 1026, A	499	26	92.9	2275	13	US-10-027-632-109912	Sequence 109912, A
C 427	26	92.9	1602	11	US-09-938-842A-1026	Sequence 1026, A	500	26	92.9	2275	15	US-10-027-632-109912	Sequence 109912, A
C 428	26	92.9	1602	11	US-10-282-122A-34610	Sequence 34610, A	501	26	92.9	2311	16	US-10-351-334-123	Sequence 123, A
C 429	26	92.9	1605	9	US-09-833-381-1807	Sequence 1807, A	502	26	92.9	2340	17	US-10-781-014-593	Sequence 593, A
C 430	26	92.9	1609	15	US-10-094-749-1505	Sequence 1505, A	503	26	92.9	2385	15	US-10-310-154-201	Sequence 201, A
C 431	26	92.9	1621	16	US-10-424-599-17136	Sequence 17136, A	504	26	92.9	2406	16	US-10-108-260A-239	Sequence 239, A
C 432	26	92.9	1626	15	US-10-172-118-1495	Sequence 1495, A	505	26	92.9	2422	17	US-10-398-036-17	Sequence 17, A
C 433	26	92.9	1626	16	US-10-342-887-1495	Sequence 1495, A	506	26	92.9	2427	15	US-10-369-493-38833	Sequence 38833, A
C 434	26	92.9	1634	15	US-10-225-066A-1017	Sequence 1017, A	507	26	92.9	2430	15	US-10-369-493-35403	Sequence 35403, A
C 435	26	92.9	1634	16	US-10-374-780A-2711	Sequence 2711, A	508	26	92.9	2463	17	US-10-437-963-37934	Sequence 37934, A

C 509	26	92.9	2475	17	US-10-437-963-68897	Sequence 68897, A	582	26	92.9	3884	14	US-10-123-108-145	Sequence 145, App
C 510	26	92.9	2500	15	US-10-310-154-194	Sequence 194, App	583	26	92.9	3884	14	US-10-123-236-145	Sequence 145, App
C 511	26	92.9	2523	15	US-10-246-785-16	Sequence 16, Appl	584	26	92.9	3884	14	US-10-123-261-145	Sequence 145, App
C 512	26	92.9	2529	15	US-10-246-785-13	Sequence 13, Appl	585	26	92.9	3884	14	US-10-140-921-145	Sequence 145, App
C 513	26	92.9	2541	15	US-10-369-493-26907	Sequence 26907, A	586	26	92.9	3884	14	US-10-140-928-145	Sequence 145, App
C 514	26	92.9	2579	10	US-09-361-652-5	Sequence 5, Appli	587	26	92.9	3884	14	US-10-121-045-145	Sequence 145, App
C 515	26	92.9	2579	10	US-09-927-315-5	Sequence 5, Appli	588	26	92.9	3884	14	US-10-123-292-145	Sequence 145, App
C 516	26	92.9	2579	15	US-10-190-417-5	Sequence 339, App	589	26	92.9	3884	14	US-10-123-903-145	Sequence 145, App
C 517	26	92.9	2595	16	US-10-282-122A-32967	Sequence 32967, A	590	26	92.9	3884	14	US-10-124-819-145	Sequence 145, App
C 518	26	92.9	2651	10	US-10-094-749-739	Sequence 739, App	591	26	92.9	3884	14	US-10-124-822-145	Sequence 145, App
C 519	26	92.9	2771	10	US-09-361-652-4	Sequence 4, Appli	592	26	92.9	3884	14	US-10-140-925-145	Sequence 145, App
C 520	26	92.9	2771	10	US-09-927-315-4	Sequence 4, Appli	593	26	92.9	3884	14	US-10-140-928-145	Sequence 145, App
C 521	26	92.9	2771	15	US-10-150-417-4	Sequence 4, Appli	594	26	92.9	3884	14	US-10-124-824-145	Sequence 145, App
C 522	26	92.9	2813	9	US-09-070-927A-459	Sequence 459, App	595	26	92.9	3884	14	US-10-127-825A-145	Sequence 145, App
C 523	26	92.9	2831	10	US-09-933-261-3	Sequence 3, Appli	596	26	92.9	3884	14	US-10-127-829A-145	Sequence 145, App
C 524	26	92.9	2831	14	US-10-256-702-3	Sequence 3, Appli	597	26	92.9	3884	14	US-10-127-835A-145	Sequence 145, App
C 525	26	92.9	2860	16	US-10-087-684-1	Sequence 1, Appli	598	26	92.9	3884	14	US-10-127-839A-145	Sequence 145, App
C 526	26	92.9	2860	16	US-10-087-684-3	Sequence 3, Appli	599	26	92.9	3884	14	US-10-127-901A-145	Sequence 145, App
C 527	26	92.9	2860	16	US-10-218-779-1	Sequence 1, Appli	600	26	92.9	3884	14	US-10-128-693A-145	Sequence 145, App
C 528	26	92.9	2860	16	US-10-218-779-3	Sequence 3, Appli	601	26	92.9	3884	14	US-10-131-813A-145	Sequence 145, App
C 529	26	92.9	2885	10	US-09-849-602-3	Sequence 3, Appli	602	26	92.9	3884	14	US-10-131-818A-145	Sequence 145, App
C 530	26	92.9	2893	17	US-10-437-963-99847	Sequence 99847, A	603	26	92.9	3884	14	US-10-131-823A-145	Sequence 145, App
C 531	26	92.9	2895	16	US-10-037-417-37	Sequence 37, Appl	604	26	92.9	3884	14	US-10-131-824A-145	Sequence 145, App
C 532	26	92.9	2915	15	US-10-104-047-1033	Sequence 1033, Ap	605	26	92.9	3884	14	US-10-131-830A-145	Sequence 145, App
C 533	26	92.9	2995	11	US-09-972-211-55	Sequence 55, Appl	606	26	92.9	3884	14	US-10-131-837A-145	Sequence 145, App
C 534	26	92.9	3000	16	US-10-096-625-55	Sequence 55, Appl	607	26	92.9	3884	14	US-10-137-872A-145	Sequence 145, App
C 535	26	92.9	3000	16	US-10-398-221-3555	Sequence 3555, Ap	608	26	92.9	3884	14	US-10-147-500-145	Sequence 145, App
C 536	26	92.9	3001	15	US-10-147-603-149	Sequence 149, App	609	26	92.9	3884	14	US-10-147-502-145	Sequence 145, App
C 537	26	92.9	3045	15	US-10-156-761-2083	Sequence 2083, Ap	610	26	92.9	3884	14	US-10-147-515-145	Sequence 145, App
C 538	26	92.9	3048	13	US-10-027-632-259317	Sequence 259317, App	611	26	92.9	3884	14	US-10-147-517-145	Sequence 145, App
C 539	26	92.9	3048	13	US-10-027-632-259318	Sequence 259318, App	612	26	92.9	3884	14	US-10-147-526-145	Sequence 145, App
C 540	26	92.9	3048	15	US-10-027-632-259317	Sequence 259317, App	613	26	92.9	3884	14	US-10-147-527-145	Sequence 145, App
C 541	26	92.9	3048	15	US-10-027-632-259318	Sequence 259318, App	614	26	92.9	3884	14	US-10-121-041-145	Sequence 145, App
C 542	26	92.9	3109	9	US-09-303-232-5	Sequence 5, Appli	615	26	92.9	3884	14	US-10-121-043-145	Sequence 145, App
C 543	26	92.9	3131	16	US-10-424-599-125726	Sequence 125726, App	616	26	92.9	3884	14	US-10-121-047-145	Sequence 145, App
C 544	26	92.9	3155	9	US-09-782-378A-15	Sequence 15, Appl	617	26	92.9	3884	14	US-10-123-215-145	Sequence 145, App
C 545	26	92.9	3177	16	US-10-282-122A-36107	Sequence 36107, A	618	26	92.9	3884	14	US-10-123-902-145	Sequence 145, App
C 546	26	92.9	3177	16	US-10-282-122A-38314	Sequence 38314, A	619	26	92.9	3884	14	US-10-123-908-145	Sequence 145, App
C 547	26	92.9	3180	16	US-10-282-122A-20891	Sequence 20891, A	620	26	92.9	3884	14	US-10-123-909-145	Sequence 145, App
C 548	26	92.9	3183	16	US-10-282-122A-21468	Sequence 21468, A	621	26	92.9	3884	14	US-10-123-910-145	Sequence 145, App
C 549	26	92.9	3189	15	US-10-369-493-41018	Sequence 41018, A	622	26	92.9	3884	14	US-10-124-813-145	Sequence 145, App
C 550	26	92.9	3211	13	US-10-027-632-114141	Sequence 114141, A	623	26	92.9	3884	14	US-10-124-817-145	Sequence 145, App
C 551	26	92.9	3211	15	US-10-027-632-114141	Sequence 114141, A	624	26	92.9	3884	14	US-10-125-922-145	Sequence 145, App
C 552	26	92.9	3314	15	US-10-104-047-865	Sequence 865, App	625	26	92.9	3884	14	US-10-125-924-145	Sequence 145, App
C 553	26	92.9	3369	10	US-09-800-187-3	Sequence 3, Appli	626	26	92.9	3884	14	US-10-140-860-145	Sequence 145, App
C 554	26	92.9	3400	16	US-10-287-326-395	Sequence 395, App	627	26	92.9	3884	14	US-10-142-417-145	Sequence 145, App
C 555	26	92.9	3485	9	US-09-816-828-18	Sequence 18, Appl	628	26	92.9	3884	14	US-10-147-519-145	Sequence 145, App
C 556	26	92.9	3501	15	US-10-295-027-1123	Sequence 1123, Ap	629	26	92.9	3884	14	US-10-157-782-145	Sequence 145, App
C 557	26	92.9	3501	17	US-10-465-498-14	Sequence 14, Appl	630	26	92.9	3884	14	US-10-152-395-145	Sequence 145, App
C 558	26	92.9	3531	13	US-10-002-600-4	Sequence 4, Appli	631	26	92.9	3884	14	US-10-125-926A-145	Sequence 145, App
C 559	26	92.9	3571	16	US-10-302-172-411	Sequence 411, App	632	26	92.9	3884	14	US-10-125-930A-145	Sequence 145, App
C 560	26	92.9	3611	13	US-10-194-163-877	Sequence 877, App	633	26	92.9	3884	14	US-10-127-831A-145	Sequence 145, App
C 561	26	92.9	3622	13	US-10-002-600-5	Sequence 5, Appli	634	26	92.9	3884	14	US-10-127-837A-145	Sequence 145, App
C 562	26	92.9	3656	13	US-10-167-264-1	Sequence 1, Appli	635	26	92.9	3884	14	US-10-127-838B-145	Sequence 145, App
C 563	26	92.9	3884	14	US-10-028-072-145	Sequence 145, App	636	26	92.9	3884	14	US-10-127-843A-145	Sequence 145, App
C 564	26	92.9	3884	14	US-10-140-808-145	Sequence 145, App	637	26	92.9	3884	14	US-10-127-843A-145	Sequence 145, App
C 565	26	92.9	3884	14	US-10-121-049-145	Sequence 145, App	638	26	92.9	3884	14	US-10-127-845A-145	Sequence 145, App
C 566	26	92.9	3884	14	US-10-123-904-145	Sequence 145, App	639	26	92.9	3884	14	US-10-127-846A-145	Sequence 145, App
C 567	26	92.9	3884	14	US-10-140-470-145	Sequence 145, App	640	26	92.9	3884	14	US-10-127-848A-145	Sequence 145, App
C 568	26	92.9	3884	14	US-10-175-746-145	Sequence 145, App	641	26	92.9	3884	14	US-10-127-849A-145	Sequence 145, App
C 569	26	92.9	3884	14	US-10-176-918-145	Sequence 145, App	642	26	92.9	3884	14	US-10-127-850A-145	Sequence 145, App
C 570	26	92.9	3884	14	US-10-176-921-145	Sequence 145, App	643	26	92.9	3884	14	US-10-127-851A-145	Sequence 145, App
C 571	26	92.9	3884	14	US-10-137-865-145	Sequence 145, App	644	26	92.9	3884	14	US-10-128-684A-145	Sequence 145, App
C 572	26	92.9	3884	14	US-10-140-474-145	Sequence 145, App	645	26	92.9	3884	14	US-10-128-686A-145	Sequence 145, App
C 573	26	92.9	3884	14	US-10-142-431-145	Sequence 145, App	646	26	92.9	3884	14	US-10-128-690A-145	Sequence 145, App
C 574	26	92.9	3884	14	US-10-143-114-145	Sequence 145, App	647	26	92.9	3884	14	US-10-128-691A-145	Sequence 145, App
C 575	26	92.9	3884	14	US-10-140-002-145	Sequence 145, App	648	26	92.9	3884	14	US-10-131-819A-145	Sequence 145, App
C 576	26	92.9	3884	14	US-10-142-419-145	Sequence 145, App	649	26	92.9	3884	14	US-10-131-829A-145	Sequence 145, App
C 577	26	92.9	3884	14	US-10-123-262-145	Sequence 145, App	650	26	92.9	3884	14	US-10-131-836A-145	Sequence 145, App
C 578	26	92.9	3884	14	US-10-142-423-145	Sequence 145, App	651	26	92.9	3884	14	US-10-146-729-145	Sequence 145, App
C 579	26	92.9	3884	14	US-10-121-050-145	Sequence 145, App	652	26	92.9	3884	14	US-10-146-791-145	Sequence 145, App
C 580	26	92.9	3884	14	US-10-141-755-145	Sequence 145, App	653	26	92.9	3884	14	US-10-147-484-145	Sequence 145, App
C 581	26	92.9	3884	14	US-10-143-032-145	Sequence 145, App	654	26	92.9	3884	14	US-10-147-508-145	Sequence 145, App

655	26	92.9	3884	14	US-10-147-512-145	Sequence 145, App	728	26	92.9	3884	15	US-10-145-870-145	Sequence 145, App
656	26	92.9	3884	14	US-10-175-735-145	Sequence 145, App	729	26	92.9	3884	15	US-10-145-876-145	Sequence 145, App
657	26	92.9	3884	14	US-10-121-040-145	Sequence 145, App	730	26	92.9	3884	15	US-10-145-859-145	Sequence 145, App
658	26	92.9	3884	14	US-10-121-056-145	Sequence 145, App	731	26	92.9	3884	15	US-10-146-724-145	Sequence 145, App
659	26	92.9	3884	14	US-10-121-061-145	Sequence 145, App	732	26	92.9	3884	15	US-10-146-725-145	Sequence 145, App
660	26	92.9	3884	14	US-10-123-235-145	Sequence 145, App	733	26	92.9	3884	15	US-10-146-795-145	Sequence 145, App
661	26	92.9	3884	14	US-10-124-818-145	Sequence 145, App	734	26	92.9	3884	15	US-10-147-495-145	Sequence 145, App
662	26	92.9	3884	14	US-10-137-868-145	Sequence 145, App	735	26	92.9	3884	15	US-10-147-501-145	Sequence 145, App
663	26	92.9	3884	14	US-10-147-492-145	Sequence 145, App	736	26	92.9	3884	15	US-10-147-504-145	Sequence 145, App
664	26	92.9	3884	14	US-10-158-782-145	Sequence 145, App	737	26	92.9	3884	15	US-10-147-506-145	Sequence 145, App
665	26	92.9	3884	14	US-10-123-905-145	Sequence 145, App	738	26	92.9	3884	15	US-10-147-509-145	Sequence 145, App
666	26	92.9	3884	14	US-10-123-907-145	Sequence 145, App	739	26	92.9	3884	15	US-10-147-510-145	Sequence 145, App
667	26	92.9	3884	14	US-10-124-815-145	Sequence 145, App	740	26	92.9	3884	15	US-10-147-511-145	Sequence 145, App
668	26	92.9	3884	14	US-10-125-921A-145	Sequence 145, App	741	26	92.9	3884	15	US-10-147-529-145	Sequence 145, App
669	26	92.9	3884	14	US-10-125-928A-145	Sequence 145, App	742	26	92.9	3884	15	US-10-152-397-145	Sequence 145, App
670	26	92.9	3884	14	US-10-127-821A-145	Sequence 145, App	743	26	92.9	3884	15	US-10-153-586-145	Sequence 145, App
671	26	92.9	3884	14	US-10-127-822A-145	Sequence 145, App	744	26	92.9	3884	15	US-10-158-786-145	Sequence 145, App
672	26	92.9	3884	14	US-10-127-824A-145	Sequence 145, App	745	26	92.9	3884	15	US-10-137-870-145	Sequence 145, App
673	26	92.9	3884	14	US-10-127-826A-145	Sequence 145, App	746	26	92.9	3884	15	US-10-140-018-145	Sequence 145, App
674	26	92.9	3884	14	US-10-127-827A-145	Sequence 145, App	747	26	92.9	3884	15	US-10-140-021-145	Sequence 145, App
675	26	92.9	3884	14	US-10-127-828A-145	Sequence 145, App	748	26	92.9	3884	15	US-10-140-471-145	Sequence 145, App
676	26	92.9	3884	14	US-10-127-830A-145	Sequence 145, App	749	26	92.9	3884	15	US-10-140-922-145	Sequence 145, App
677	26	92.9	3884	14	US-10-127-832A-145	Sequence 145, App	750	26	92.9	3884	15	US-10-145-631-145	Sequence 145, App
678	26	92.9	3884	14	US-10-127-834A-145	Sequence 145, App	751	26	92.9	3884	15	US-10-145-633-145	Sequence 145, App
679	26	92.9	3884	14	US-10-127-834A-145	Sequence 145, App	752	26	92.9	3884	15	US-10-158-783-145	Sequence 145, App
680	26	92.9	3884	14	US-10-127-836A-145	Sequence 145, App	753	26	92.9	3884	15	US-10-140-274-145	Sequence 145, App
681	26	92.9	3884	14	US-10-127-841A-145	Sequence 145, App	754	26	92.9	3884	15	US-10-140-019-145	Sequence 145, App
682	26	92.9	3884	14	US-10-128-687A-145	Sequence 145, App	755	26	92.9	3884	15	US-10-140-022-145	Sequence 145, App
683	26	92.9	3884	14	US-10-128-687A-145	Sequence 145, App	756	26	92.9	3884	15	US-10-140-861-145	Sequence 145, App
684	26	92.9	3884	14	US-10-128-688A-145	Sequence 145, App	757	26	92.9	3884	15	US-10-140-862-145	Sequence 145, App
685	26	92.9	3884	14	US-10-128-689A-145	Sequence 145, App	758	26	92.9	3884	15	US-10-141-697-145	Sequence 145, App
686	26	92.9	3884	14	US-10-128-694A-145	Sequence 145, App	759	26	92.9	3884	15	US-10-141-700-145	Sequence 145, App
687	26	92.9	3884	14	US-10-131-825A-145	Sequence 145, App	760	26	92.9	3884	15	US-10-141-705-145	Sequence 145, App
688	26	92.9	3884	14	US-10-230-417-145	Sequence 145, App	761	26	92.9	3884	15	US-10-141-753-145	Sequence 145, App
689	26	92.9	3884	14	US-10-131-815A-145	Sequence 145, App	762	26	92.9	3884	15	US-10-141-758-145	Sequence 145, App
690	26	92.9	3884	14	US-10-131-817A-145	Sequence 145, App	763	26	92.9	3884	15	US-10-142-418-145	Sequence 145, App
691	26	92.9	3884	14	US-10-131-821A-145	Sequence 145, App	764	26	92.9	3884	15	US-10-142-420-145	Sequence 145, App
692	26	92.9	3884	14	US-10-131-822A-145	Sequence 145, App	765	26	92.9	3884	15	US-10-142-422-145	Sequence 145, App
693	26	92.9	3884	14	US-10-131-828A-145	Sequence 145, App	766	26	92.9	3884	15	US-10-142-427-145	Sequence 145, App
694	26	92.9	3884	14	US-10-131-835A-145	Sequence 145, App	767	26	92.9	3884	15	US-10-142-760-145	Sequence 145, App
695	26	92.9	3884	14	US-10-137-864A-145	Sequence 145, App	768	26	92.9	3884	15	US-10-145-821-145	Sequence 145, App
696	26	92.9	3884	14	US-10-137-869A-145	Sequence 145, App	769	26	92.9	3884	15	US-10-152-531-145	Sequence 145, App
697	26	92.9	3884	14	US-10-147-523-145	Sequence 145, App	770	26	92.9	3884	15	US-10-127-840A-145	Sequence 145, App
698	26	92.9	3884	14	US-10-158-785-145	Sequence 145, App	771	26	92.9	3884	15	US-10-142-424-145	Sequence 145, App
699	26	92.9	3884	14	US-10-121-051-145	Sequence 145, App	772	26	92.9	3884	15	US-10-142-761-145	Sequence 145, App
700	26	92.9	3884	14	US-10-121-042-145	Sequence 145, App	773	26	92.9	3884	15	US-10-142-763-145	Sequence 145, App
701	26	92.9	3884	15	US-10-123-913-145	Sequence 145, App	774	26	92.9	3884	15	US-10-142-765-145	Sequence 145, App
702	26	92.9	3884	15	US-10-192-007-145	Sequence 145, App	775	26	92.9	3884	15	US-10-142-887-145	Sequence 145, App
703	26	92.9	3884	15	US-10-194-359-145	Sequence 145, App	776	26	92.9	3884	15	US-10-142-888-145	Sequence 145, App
704	26	92.9	3884	15	US-10-175-590-145	Sequence 145, App	777	26	92.9	3884	15	US-10-145-015-145	Sequence 145, App
705	26	92.9	3884	15	US-10-137-866-145	Sequence 145, App	778	26	92.9	3884	15	US-10-145-090-145	Sequence 145, App
706	26	92.9	3884	15	US-10-146-726-145	Sequence 145, App	779	26	92.9	3884	15	US-10-143-116-145	Sequence 145, App
707	26	92.9	3884	15	US-10-146-727-145	Sequence 145, App	780	26	92.9	3884	15	US-10-143-117-145	Sequence 145, App
708	26	92.9	3884	15	US-10-146-728-145	Sequence 145, App	781	26	92.9	3884	15	US-10-144-957-145	Sequence 145, App
709	26	92.9	3884	15	US-10-146-788-145	Sequence 145, App	782	26	92.9	3884	15	US-10-144-992-145	Sequence 145, App
710	26	92.9	3884	15	US-10-152-380-145	Sequence 145, App	783	26	92.9	3884	15	US-10-145-015-145	Sequence 145, App
711	26	92.9	3884	15	US-10-153-934-145	Sequence 145, App	784	26	92.9	3884	15	US-10-145-091-145	Sequence 145, App
712	26	92.9	3884	15	US-10-140-807-145	Sequence 145, App	785	26	92.9	3884	15	US-10-145-629-145	Sequence 145, App
713	26	92.9	3884	15	US-10-140-924-145	Sequence 145, App	786	26	92.9	3884	15	US-10-145-630-145	Sequence 145, App
714	26	92.9	3884	15	US-10-140-926-145	Sequence 145, App	787	26	92.9	3884	15	US-10-145-747-145	Sequence 145, App
715	26	92.9	3884	15	US-10-141-698-145	Sequence 145, App	788	26	92.9	3884	15	US-10-145-752-145	Sequence 145, App
716	26	92.9	3884	15	US-10-141-702-145	Sequence 145, App	789	26	92.9	3884	15	US-10-145-754-145	Sequence 145, App
717	26	92.9	3884	15	US-10-141-704-145	Sequence 145, App	790	26	92.9	3884	15	US-10-145-755-145	Sequence 145, App
718	26	92.9	3884	15	US-10-142-421-145	Sequence 145, App	791	26	92.9	3884	15	US-10-145-818-145	Sequence 145, App
719	26	92.9	3884	15	US-10-142-432-145	Sequence 145, App	792	26	92.9	3884	15	US-10-145-820-145	Sequence 145, App
720	26	92.9	3884	15	US-10-142-767-145	Sequence 145, App	793	26	92.9	3884	15	US-10-145-872-145	Sequence 145, App
721	26	92.9	3884	15	US-10-143-033-145	Sequence 145, App	794	26	92.9	3884	15	US-10-145-873-145	Sequence 145, App
722	26	92.9	3884	15	US-10-144-994-145	Sequence 145, App	795	26	92.9	3884	15	US-10-147-481-145	Sequence 145, App
723	26	92.9	3884	15	US-10-145-628-145	Sequence 145, App	796	26	92.9	3884	15	US-10-147-482-145	Sequence 145, App
724	26	92.9	3884	15	US-10-145-746-145	Sequence 145, App	797	26	92.9	3884	15	US-10-147-503-145	Sequence 145, App
725	26	92.9	3884	15	US-10-145-748-145	Sequence 145, App	798	26	92.9	3884	15	US-10-147-522-145	Sequence 145, App
726	26	92.9	3884	15	US-10-145-823-145	Sequence 145, App	799	26	92.9	3884	15	US-10-152-401-145	Sequence 145, App
727	26	92.9	3884	15	US-10-145-825-145	Sequence 145, App	800	26	92.9	3884	15	US-10-157-783-145	Sequence 145, App

801	26	92.9	3884	15	US-10-158-792-145	Sequence 145, App
802	26	92.9	3884	15	US-10-158-462-145	Sequence 145, App
803	26	92.9	3884	15	US-10-143-035-145	Sequence 145, App
804	26	92.9	3884	15	US-10-145-751-145	Sequence 145, App
805	26	92.9	3884	15	US-10-145-822-145	Sequence 145, App
806	26	92.9	3884	15	US-10-145-824-145	Sequence 145, App
807	26	92.9	3884	15	US-10-145-867-145	Sequence 145, App
808	26	92.9	3884	15	US-10-145-869-145	Sequence 145, App
809	26	92.9	3884	15	US-10-145-875-145	Sequence 145, App
810	26	92.9	3884	15	US-10-145-877-145	Sequence 145, App
811	26	92.9	3884	15	US-10-145-958-145	Sequence 145, App
812	26	92.9	3884	15	US-10-146-787-145	Sequence 145, App
813	26	92.9	3884	15	US-10-146-790-145	Sequence 145, App
814	26	92.9	3884	15	US-10-146-793-145	Sequence 145, App
815	26	92.9	3884	15	US-10-147-480-145	Sequence 145, App
816	26	92.9	3884	15	US-10-147-485-145	Sequence 145, App
817	26	92.9	3884	15	US-10-147-486-145	Sequence 145, App
818	26	92.9	3884	15	US-10-147-487-145	Sequence 145, App
819	26	92.9	3884	15	US-10-147-490-145	Sequence 145, App
820	26	92.9	3884	15	US-10-147-494-145	Sequence 145, App
821	26	92.9	3884	15	US-10-147-498-145	Sequence 145, App
822	26	92.9	3884	15	US-10-147-514-145	Sequence 145, App
823	26	92.9	3884	15	US-10-147-524-145	Sequence 145, App
824	26	92.9	3884	15	US-10-152-379-145	Sequence 145, App
825	26	92.9	3884	15	US-10-152-394-145	Sequence 145, App
826	26	92.9	3884	15	US-10-152-406-145	Sequence 145, App
827	26	92.9	3884	15	US-10-156-847-145	Sequence 145, App
828	26	92.9	3884	15	US-10-157-778-145	Sequence 145, App
829	26	92.9	3884	15	US-10-157-799-145	Sequence 145, App
830	26	92.9	3884	15	US-10-160-504-145	Sequence 145, App
831	26	92.9	3884	15	US-10-145-634-145	Sequence 145, App
832	26	92.9	3884	15	US-10-147-520-145	Sequence 145, App
833	26	92.9	3884	15	US-10-157-781-145	Sequence 145, App
834	26	92.9	3884	15	US-10-176-989-145	Sequence 145, App
835	26	92.9	3884	15	US-10-147-491-145	Sequence 145, App
836	26	92.9	3884	15	US-10-152-378-145	Sequence 145, App
837	26	92.9	3884	15	US-10-152-382-145	Sequence 145, App
838	26	92.9	3884	15	US-10-152-383-145	Sequence 145, App
839	26	92.9	3884	15	US-10-152-384-145	Sequence 145, App
840	26	92.9	3884	15	US-10-152-387-145	Sequence 145, App
841	26	92.9	3884	15	US-10-152-389-145	Sequence 145, App
842	26	92.9	3884	15	US-10-152-390-145	Sequence 145, App
843	26	92.9	3884	15	US-10-152-392-145	Sequence 145, App
844	26	92.9	3884	15	US-10-152-403-145	Sequence 145, App
845	26	92.9	3884	15	US-10-153-756-145	Sequence 145, App
846	26	92.9	3884	15	US-10-157-779-145	Sequence 145, App
847	26	92.9	3884	15	US-10-157-784-145	Sequence 145, App
848	26	92.9	3884	15	US-10-157-797-145	Sequence 145, App
849	26	92.9	3884	15	US-10-158-491-145	Sequence 145, App
850	26	92.9	3884	15	US-10-142-762-145	Sequence 145, App
851	26	92.9	3884	15	US-10-142-764-145	Sequence 145, App
852	26	92.9	3884	15	US-10-142-766-145	Sequence 145, App
853	26	92.9	3884	15	US-10-143-625-145	Sequence 145, App
854	26	92.9	3884	15	US-10-145-627-145	Sequence 145, App
855	26	92.9	3884	15	US-10-145-960-145	Sequence 145, App
856	26	92.9	3884	15	US-10-145-962-145	Sequence 145, App
857	26	92.9	3884	15	US-10-146-789-145	Sequence 145, App
858	26	92.9	3884	15	US-10-147-483-145	Sequence 145, App
859	26	92.9	3884	15	US-10-147-496-145	Sequence 145, App
860	26	92.9	3884	15	US-10-147-505-145	Sequence 145, App
861	26	92.9	3884	15	US-10-147-516-145	Sequence 145, App
862	26	92.9	3884	15	US-10-152-398-145	Sequence 145, App
863	26	92.9	3884	15	US-10-152-398-145	Sequence 145, App
864	26	92.9	3884	15	US-10-139-980-145	Sequence 145, App
865	26	92.9	3884	15	US-10-145-750-145	Sequence 145, App
866	26	92.9	3884	15	US-10-152-373-145	Sequence 145, App
867	26	92.9	3884	15	US-10-121-044-145	Sequence 145, App
868	26	92.9	3884	15	US-10-117-505-145	Sequence 145, App
869	26	92.9	3884	15	US-10-121-055-145	Sequence 145, App
870	26	92.9	3884	15	US-10-121-057-145	Sequence 145, App
871	26	92.9	3884	15	US-10-121-058-145	Sequence 145, App
872	26	92.9	3884	15	US-10-121-060-145	Sequence 145, App
873	26	92.9	3884	15	US-10-123-109-145	Sequence 145, App
874	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
875	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
876	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
877	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
878	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
879	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
880	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
881	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
882	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
883	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
884	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
885	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
886	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
887	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
888	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
889	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
890	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
891	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
892	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
893	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
894	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
895	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
896	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
897	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
898	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
899	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
900	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
901	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
902	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
903	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
904	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
905	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
906	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
907	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
908	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
909	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
910	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
911	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
912	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
913	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
914	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
915	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
916	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
917	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
918	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
919	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
920	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
921	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
922	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
923	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
924	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
925	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
926	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
927	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
928	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
929	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
930	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
931	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
932	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
933	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
934	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
935	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
936	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
937	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
938	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
939	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
940	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
941	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
942	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
943	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
944	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
945	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
946	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
947	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
948	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
949	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
950	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
951	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
952	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
953	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
954	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
955	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
956	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
957	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
958	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
959	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
960	26	92.9	3884	15	US-10-123-154-145	Sequence 145, App
961	26	92.9	3884	15	US-10-123-154-145	

947 26 92.9 3884 15 US-10-142-884-145 Sequence 145, App
948 26 92.9 3884 15 US-10-143-027-145 Sequence 145, App
949 26 92.9 3884 15 US-10-143-115-145 Sequence 145, App
950 26 92.9 3884 15 US-10-144-956-145 Sequence 145, App
951 26 92.9 3884 15 US-10-144-958-145 Sequence 145, App
952 26 92.9 3884 15 US-10-145-632-145 Sequence 145, App
953 26 92.9 3884 15 US-10-145-749-145 Sequence 145, App
954 26 92.9 3884 15 US-10-145-753-145 Sequence 145, App
955 26 92.9 3884 15 US-10-145-871-145 Sequence 145, App
956 26 92.9 3884 15 US-10-145-878-145 Sequence 145, App
957 26 92.9 3884 15 US-10-146-794-145 Sequence 145, App
958 26 92.9 3884 15 US-10-147-489-145 Sequence 145, App
959 26 92.9 3884 15 US-10-147-507-145 Sequence 145, App
960 26 92.9 3884 15 US-10-147-535-145 Sequence 145, App
961 26 92.9 3884 15 US-10-147-537-145 Sequence 145, App
962 26 92.9 3884 15 US-10-152-376-145 Sequence 145, App
963 26 92.9 3884 15 US-10-152-381-145 Sequence 145, App
964 26 92.9 3884 15 US-10-152-400-145 Sequence 145, App
965 26 92.9 3884 15 US-10-153-585-145 Sequence 145, App
966 26 92.9 3884 15 US-10-157-780-145 Sequence 145, App
967 26 92.9 3884 15 US-10-157-800-145 Sequence 145, App
968 26 92.9 3884 15 US-10-157-801-145 Sequence 145, App
969 26 92.9 3884 15 US-10-157-802-145 Sequence 145, App
970 26 92.9 3884 15 US-10-158-784-145 Sequence 145, App
971 26 92.9 3884 15 US-10-158-789-145 Sequence 145, App
972 26 92.9 3884 15 US-10-192-011-145 Sequence 145, App
973 26 92.9 3884 15 US-10-139-963-145 Sequence 145, App
974 26 92.9 3884 15 US-10-140-020-145 Sequence 145, App
975 26 92.9 3884 15 US-10-140-023-145 Sequence 145, App
976 26 92.9 3884 15 US-10-140-809-145 Sequence 145, App
977 26 92.9 3884 15 US-10-140-865-145 Sequence 145, App
978 26 92.9 3884 15 US-10-141-701-145 Sequence 145, App
979 26 92.9 3884 15 US-10-141-754-145 Sequence 145, App
980 26 92.9 3884 15 US-10-141-760-145 Sequence 145, App
981 26 92.9 3884 15 US-10-142-425-145 Sequence 145, App
982 26 92.9 3884 15 US-10-142-430-145 Sequence 145, App
983 26 92.9 3884 15 US-10-143-113-145 Sequence 145, App
984 26 92.9 3884 15 US-10-146-730-145 Sequence 145, App
985 26 92.9 3884 15 US-10-146-792-145 Sequence 145, App
986 26 92.9 3884 15 US-10-158-791-145 Sequence 145, App
987 26 92.9 3884 15 US-10-156-843-145 Sequence 145, App
988 26 92.9 3884 15 US-10-157-786-145 Sequence 145, App
989 26 92.9 3884 15 US-10-152-405-145 Sequence 145, App
990 26 92.9 3884 15 US-10-147-528-145 Sequence 145, App
991 26 92.9 3884 16 US-10-128-692A-145 Sequence 145, App
992 26 92.9 3884 16 US-10-140-927-145 Sequence 145, App
993 26 92.9 3884 16 US-10-147-493-145 Sequence 145, App
994 26 92.9 3884 16 US-10-145-127-145 Sequence 145, App
995 26 92.9 3884 16 US-10-160-503-145 Sequence 145, App
996 26 92.9 3884 16 US-10-143-118-145 Sequence 145, App
997 26 92.9 3884 16 US-10-144-993-145 Sequence 145, App
998 26 92.9 3884 16 US-10-158-787-145 Sequence 145, App
999 26 92.9 3884 16 US-10-140-024-145 Sequence 145, App
1000 26 92.9 3884 16 US-10-147-536-145 Sequence 145, App

ALIGNMENTS

RESULT 1
US-10-242-535A-58809
; Sequence 58809, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58809
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-58809

Alignment Scores:
Pred. No.: 185 Length: 200
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-242-535A-58809 (1-200)

Oy 1 GlyTyr***ValGluGlu 6
Db 63 GGATATTCTGTAGAAGAA 80
||||| ||||| ||||| ||||| |||||

RESULT 2

US-10-085-783A-58809
; Sequence 58809, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58809
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-58809

Alignment Scores:
Pred. No.: 185 Length: 200
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-085-783A-58809 (1-200)

Oy 1 GlyTyr***ValGluGlu 6
Db 63 GGATATTCTGTAGAAGAA 80
||||| ||||| ||||| ||||| |||||

RESULT 3

US-09-864-408A-3483
; Sequence 3483, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encor

; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3483
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-864-408A-3483

Alignment Scores:
Pred. No.: 305 Length: 318
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 11 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-864-408A-3483 (1-318)

Qy 1 GlyTyr***ValGluGlu 6
Db 165 GGATATTCTGTGAAGAA 182

RESULT 4

US-09-560-863-269/c
; Sequence 269, Application US/09560863
; Patent No. US20020110809A1
; GENERAL INFORMATION:
; APPLICANT: Nehls, Michael C.
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020110809A1el Human Polynucleotides and the
; FILE REFERENCE: Polypeptides Encoded Thereby
; CURRENT APPLICATION NUMBER: US/09/560,863
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/132,408
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 1008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269
; LENGTH: 384
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(384)
; OTHER INFORMATION: n = A,T,C or G
US-09-560-863-269

Alignment Scores:
Pred. No.: 374 Length: 384
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-560-863-269 (1-384)

Qy 1 GlyTyr***valGluGlu 6
Db 328 GGGTACAGTGTGAAGAG 311

RESULT 5

US-10-767-701-30758/c
; Sequence 30758, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 30758
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 18059428
US-10-767-701-30758

Alignment Scores:
Pred. No.: 384 Length: 394
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-767-701-30758 (1-394)

Qy 1 GlyTyr***ValGluGlu 6
Db 367 GGCTACACGTCGAGGAG 350

RESULT 6

US-10-437-963-82192
; Sequence 82192, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 82192
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81646C.1
US-10-437-963-82192

Alignment Scores:
Pred. No.: 400 Length: 409
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-437-963-82192 (1-409)

Qy 1 GlyTyr***ValGluGlu 6
Db 158 GGCTACACGTCGAGGAA 175

RESULT 7

Alignment Scores:

Pred. No.: 445 Length: 452
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-783-590-8368 (1-452)

Qy 1 GlyTyr***ValGluGlu 6

Db 172 GGTATAGTGTGGAAGAG 155

RESULT 9

US-10-021-323-11049
; Sequence 11049, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 11049
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-033-Q6-N6-E12
US-10-021-323-11049

Alignment Scores:
Pred. No.: 470 Length: 475
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-021-323-11049 (1-475)

Qy 1 GlyTyr***ValGluGlu 6

Db 223 GGTATTCAGTAGAGAA 240

RESULT 10

US-10-021-323-10992/c
; Sequence 10992, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 10992
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Gossypium hirsutum

; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-033-Q6-K6-E12
US-10-021-323-10992
Alignment Scores:
Pred. No.: 471 Length: 476
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-021-323-10992 (1-476)

Qy 1 GlyTyr***ValGluGlu 6

Db 254 GGTATTCAGTAGAGAA 237

RESULT 11

US-09-974-300-4300
; Sequence 4300, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085-500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4300
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-4300

Alignment Scores:
Pred. No.: 474 Length: 479
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-974-300-4300 (1-479)

Qy 1 GlyTyr***ValGluGlu 6

Db 334 GGATACAGTGTGGAAGAG 351

RESULT 12

US-10-425-114-35722
; Sequence 35722, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128


```
; OTHER INFORMATION: Clone ID: 700975650_FLI
US-10-425-114-10779
Alignment Scores:
Pred. No.: 540 Length: 540
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-425-114-10779 (1-540)
Qy 1 GlyTyr***ValGluGlu 6
Db 333 GGGTACTCCGTGGAAGAG 350
RESULT 16
US-10-021-323-10850/c
; Sequence 10850, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 10850
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-031-Q6-N6-G5
US-10-021-323-10850
Alignment Scores:
Pred. No.: 543 Length: 543
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-021-323-10850 (1-543)
Qy 1 GlyTyr***ValGluGlu 6
Db 159 GGATATTCTGTGAAGAG 142
RESULT 17
US-10-021-323-3889/c
; Sequence 3889, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 3889
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3825-029-Q6-N6-C3
US-10-021-323-3889
Alignment Scores:
Pred. No.: 544 Length: 544
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-021-323-3889 (1-544)
Qy 1 GlyTyr***ValGluGlu 6
Db 82 GGATATTCTGTGAAGAG 65
RESULT 18
US-10-021-323-11186/c
; Sequence 11186, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 11186
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-034-Q6-N6-G5
US-10-021-323-11186
Alignment Scores:
Pred. No.: 545 Length: 545
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-021-323-11186 (1-545)
Qy 1 GlyTyr***ValGluGlu 6
Db 159 GGATATTCTGTGAAGAG 142
RESULT 19
US-10-027-632-210714/c
; Sequence 210714, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827, 129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
```

```

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 210714
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-210714
Alignment Scores:
Pred. No.: 574 Length: 572
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 13 Gaps: 0
214 GGGTACAGTGTGAGGAA 197

US-10-030-194A-5 (1-6) x US-10-027-632-210714 (1-572)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 214 GGGTACAGTGTGAGGAA 197

RESULT 20
US-10-027-632-210714/c
; Sequence 210714, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 210714
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-210714
Alignment Scores:
Pred. No.: 574 Length: 572
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 13 Gaps: 0
214 GGGTACAGTGTGAGGAA 197

```

```

Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-027-632-210714 (1-572)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 214 GGGTACAGTGTGAGGAA 197

RESULT 21
US-10-027-632-194951/c
; Sequence 194951, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194951
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-194951
Alignment Scores:
Pred. No.: 635 Length: 628
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 13 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-027-632-194951 (1-628)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 504 GGTACAGTGTGAGGAA 487

RESULT 22
US-10-027-632-194951/c
; Sequence 194951, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29

```

RESULT 24
US-10-027-632-8887
; Sequence 8887, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and

```

; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8888
; LENGTH: 659
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-8888

```



```
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-027-632-8888 (1-659)

Qy 1 GlyTyr***ValGluGlu 6
|||
Db 419 GGATACCTCAGTTGAGGAA 436

RESULT 29
US-10-027-632-8889
; Sequence 8889, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8889
; LENGTH: 659
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-8889

Alignment Scores:
Pred. No.: 669 Length: 659
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-027-632-8889 (1-659)

Qy 1 GlyTyr***ValGluGlu 6
|||
Db 419 GGATACCTCAGTTGAGGAA 436

RESULT 30
US-10-027-632-226071/c
; Sequence 226071, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8889
; LENGTH: 659
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-8889

Alignment Scores:
Pred. No.: 669 Length: 659
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0
```

```
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 226071
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-226071

Alignment Scores:
Pred. No.: 674 Length: 664
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 13 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-027-632-226071 (1-664)

Qy 1 GlyTyr***ValGluGlu 6
|||
Db 584 GGCTATTCTGTGGAGGAG 567

RESULT 31
US-10-027-632-226071/c
; Sequence 226071, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 226071
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-226071

Alignment Scores:
Pred. No.: 674 Length: 664
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-027-632-226071 (1-664)
```

```
Qy 1 GlyTyr***ValGluGlu 6
Db 584 GGCTATTCTGTGGAGGAG 567

RESULT 32
US-09-764-887-485/c
; Sequence 485, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P113
; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 485
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-887-485

Alignment Scores:
Pred. No.: 766 Length: 747
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194a-5 (1-6) x US-09-764-887-485 (1-747)

Qy 1 GlyTyr***ValGluGlu 6
Db 203 GGTTATTCTGTGAAGAA 186

RESULT 33
US-10-073-961-485/c
; Sequence 485, Application US/10073961
; Publication No. US20030077602A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P113C1
; CURRENT APPLICATION NUMBER: US/10/073,961
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/764,887
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
```

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14

; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Alignment Scores:

Pred. No.:	766	Length:	747
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	14	Gaps:	0

US-10-030-194A-5 (1-6) x US-10-073-961-485 (1-747)

Oy 1 GlyTyr***ValGluGlu 6
||||| ||||| ||||| |||||
Db 203 GGTATTCTGTGGAAGAA 186

RESULT 34

US-10-424-599-51314/c
; Sequence 51314, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 51314
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_17350C.1
US-10-424-599-51314

Alignment Scores:

Pred. No.:	779	Length:	759
Score:	27.00	Matches:	5
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	1
Query Match:	96.43%	Indels:	0
DB:	16	Gaps:	0

US-10-030-194A-5 (1-6) x US-10-424-599-51314 (1-759)

Oy 1 GlyTyr***ValGluGlu 6
||||| ||||| ||||| |||||
Db 35 GGATATGCCGTAGAGAG 18

RESULT 35

```
US-10-282-122A-6833/c
; Sequence 6833, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6833
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-282-122A-6833

Alignment Scores:
Pred. No.: 862 Length: 834
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-282-122A-6833 (1-834)
Qy 1 GlyTyr***ValGluGlu 6
Db 336 GGTACGCTGTCGAGGAA 319

RESULT 36
US-10-369-493-33361
; Sequence 33361, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33361
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Desulfitobacterium hafniese
US-10-369-493-33361

Alignment Scores:
Pred. No.: 866 Length: 837
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-369-493-33361 (1-837)
Qy 1 GlyTyr***ValGluGlu 6
Db 208 GGATATAGCGTAGAAGAA 225

RESULT 37
US-10-369-493-23809
; Sequence 23809, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23809
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Aquifex aeolicus
US-10-369-493-23809

Alignment Scores:
Pred. No.: 983 Length: 942
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-369-493-23809 (1-942)
Qy 1 GlyTyr***ValGluGlu 6
Db 736 GGATACAGCGTTGAGGAA 753

RESULT 38
US-10-632-171842/c
; Sequence 171842, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```

; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 171842
; LENGTH: 947
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-171842

Alignment Scores:
Pred. No.: 989 Length: 947
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 13 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-027-632-171842 (1-947)
Qy 1 GlyTyr***ValGluGlu 6
Db 741 GGATACAGTGTGAAGAA 724

RESULT 39
US-10-027-632-171842/c
; Sequence 171842, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 171842
; LENGTH: 947
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-171842

```

```

Alignment Scores:
Pred. No.: 989 Length: 947
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-027-632-171842 (1-947)
Qy 1 GlyTyr***ValGluGlu 6
Db 741 GGATACAGTGTGAAGAA 724

RESULT 40
US-10-369-493-45793
; Sequence 45793, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 45793
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-45793

Alignment Scores:
Pred. No.: 1,01e+03 Length: 969
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-369-493-45793 (1-969)
Qy 1 GlyTyr***ValGluGlu 6
Db 678 GGATATTCTGTGAAGAA 695

RESULT 41
US-10-767-795-1815/c
; Sequence 1815, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 1815
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C670_2
US-10-767-795-1815

```

```

RESULT 43
US-09-804-291-302/c
; Sequence 302, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:
; APPLICANT: ZOZULA, SERGEY
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: P 0278005
; CURRENT APPLICATION NUMBER: US/09/804,291
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/188,914
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/192,033
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/198,474
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/199,335
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: 60/207,702
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR FILING DATE: 2000-06-23

```

APPLICANT: PATTERSON, Chandra

APPLICANT: YAO, Monique G.
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: THORNTON, Michael
APPLICANT: LU, Yan
APPLICANT: KALLICK, Deborah A.
APPLICANT: GANDHI, Ameena R.
APPLICANT: AU-YOUNG, Janice
TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
FILE REFERENCE: SF-0781 PCT
CURRENT APPLICATION NUMBER: US/10/297,021
CURRENT FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: 60/206,222; 60/207,476; 60/208,834; 60/208,861; 60/209,868
PRIOR FILING DATE: 2000-05-22; 2000-05-25; 2000-06-02; 2000-06-02; 2000-06-07
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PERL Program
SEQ ID NO 44
LENGTH: 987
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040023294A1 7476117CB1
US-10-297-021-44

Alignment Scores:
Pred. No.: 1.03e+03 Length: 987
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-297-021-44 (1-987)

Qy 1 GlyTyr***ValGluGlu 6
Db 792 GGGTACACAGTAGAGGAA 775

RESULT 46
US-10-343-650A-361/c
Sequence 361, Application US/10343650A
Publication No. US20040067499A1
GENERAL INFORMATION:
APPLICANT: HAGA, TATSUYA
TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
FILE REFERENCE: 31671-186347
CURRENT APPLICATION NUMBER: US/10/343,650A
CURRENT FILING DATE: 2003-07-21
PRIOR APPLICATION NUMBER: JP 2000/237818
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: JP 2001/34434
PRIOR FILING DATE: 2001-02-13
NUMBER OF SEQ ID NOS: 694
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 361
LENGTH: 987
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(987)
US-10-343-650A-361

Alignment Scores:
Pred. No.: 1.03e+03 Length: 987
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-343-650A-361 (1-987)

Qy 1 GlyTyr***ValGluGlu 6
Db 792 GGGTACACAGTAGAGGAA 775
RESULT 47
US-10-024-399-31/c
Sequence 31, Application US/10024399
Publication No. US20030100491A1
GENERAL INFORMATION:
APPLICANT: Padigar, Muralidhara
APPLICANT: Kekuda, Ramesh
APPLICANT: Colman, Steven D.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Ballinger, Robert A.
APPLICANT: Vernet, Corine A.M.
APPLICANT: Li, Li
APPLICANT: Shenoy, Suresh G.
APPLICANT: Casman, Stacie J.
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-224AE
CURRENT APPLICATION NUMBER: US/10/024,399
CURRENT FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: 60/256,635
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/259,743
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: 60/299,327
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: 60/261,498
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/263,689
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/267,464
PRIOR FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: 60/271,021
PRIOR FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: 60/275,946
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/278,150
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: 60/285,718
PRIOR FILING DATE: 2001-04-23
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 1002
TYPE: DNA
ORGANISM: Homo sapiens
US-10-024-399-31

Alignment Scores:
Pred. No.: 1.05e+03 Length: 1002
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-024-399-31 (1-1002)

Qy 1 GlyTyr***ValGluGlu 6
Db 798 GGGTACACAGTAGAGGAA 781

RESULT 48
US-10-437-963-73808
Sequence 73808, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua


```
Qy 1 GlyTyr***ValGluGlu 6
Db 964 GGCTACGCGGTGGAGGAA 981

RESULT 51
US-10-369-493-40154
; Sequence 40154, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 40154
; LENGTH: 1099
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-369-493-40154

Alignment Scores:
Pred. No.: 1.16e+03 Length: 1099
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-369-493-40154 (1-1099)

Qy 1 GlyTyr***ValGluGlu 6
Db 865 GGATACGCGGTAGAGAA 882

RESULT 52
US-10-767-795-1814/c
; Sequence 1814, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 1814
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C670_3
US-10-767-795-1814

Alignment Scores:
Pred. No.: 1.17e+03 Length: 1110
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-767-795-1814 (1-1110)

Qy 1 GlyTyr***ValGluGlu 6
Db 487 GGTATTTCAGTAGAGAA 470

RESULT 53
US-10-260-238-779
; Sequence 779, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 779
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (127)..(127)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (129)..(129)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-779

Alignment Scores:
Pred. No.: 1.36e+03 Length: 1269
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-260-238-779 (1-1269)

Qy 1 GlyTyr***ValGluGlu 6
Db 313 GGATACGCGGTGGAGGAG 330

RESULT 54
US-10-437-963-15300
; Sequence 15300, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
```

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15300
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21156C.1
US-10-437-963-15300

Alignment Scores:
Pred. No.: 1.36e+03 Length: 1269
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-437-963-15300 (1-1269)

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 313 GGATACCTCGGTGGAGAG 330

RESULT 55

US-10-425-114-20541
; Sequence 20541, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20541
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3242-059-D4_FLI
US-10-425-114-20541

Alignment Scores:
Pred. No.: 1.37e+03 Length: 1282
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-425-114-20541 (1-1282)

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 1075 GGGTACTCCGTGGAAGAG 1092

RESULT 56

US-10-437-963-25616/c
; Sequence 25616, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 25616
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_30486C.1
US-10-437-963-25616

Alignment Scores:
Pred. No.: 1.46e+03 Length: 1362
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-437-963-25616 (1-1362)

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 504 GGCTATCTAGTGAAGAA 487

RESULT 57

US-10-369-493-36451
; Sequence 36451, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 36451
; LENGTH: 1378
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-10-369-493-36451

Alignment Scores:
Pred. No.: 1.48e+03 Length: 1378
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-369-493-36451 (1-1378)

Qy 1 GlyTyr***ValGluGlu 6
|||||
Db 275 GGATATACGTCGAAGAG 292

```

; LOCATION: (201)..(1187)
US-10-292-798-255

Alignment Scores:
Pred. No.: 1.49e+03 Length: 1387
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-292-798-255 (1-1387)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 992 GGGTACACAGTAGAGAA 975

RESULT 60
US-10-354-437-31
; Sequence 31, Application US/10354437
; Publication No. US20040023257A1
; GENERAL INFORMATION:
; APPLICANT: Barton, Nelson R.
; APPLICANT: Weiner, David Paul
; APPLICANT: Greenberg, William
; APPLICANT: Luu, Samantha
; APPLICANT: Chang, Kristine
; APPLICANT: Waters, Elizabeth
; TITLE OF INVENTION: ENZYMES HAVING SECONDARY AMIDASES ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-106001
; CURRENT APPLICATION NUMBER: US/10/354,437
; PRIOR FILING DATE: 2003-01-20
; PRIOR APPLICATION NUMBER: US 60/352,895
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 1425
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
US-10-354-437-31

Alignment Scores:
Pred. No.: 1.54e+03 Length: 1425
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-354-437-31 (1-1425)

Qy 1 GlyTyr***ValGluGlu 6
||||| |||||||
Db 877 GGCATCGCGTGAGGAG 894

RESULT 61
US-09-815-242-4536
; Sequence 4536, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in

```

```
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4536
; LENGTH: 1635
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4536
```

```
Alignment Scores:
Pred. No.: 1.78e+03 Length: 1635
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0
```

US-10-030-194A-5 (1-6) x US-09-815-242-4536 (1-1635)

```
Oy 1 GlyTyr**ValGluGlu 6
      ||||| ||||| ||||| |||||
Db 805 GGATACCTCTGTAGAAGAA 822
```

```
RESULT 62
US-09-815-242-8100
; Sequence 8100, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8100
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1662)
US-09-815-242-8100

Alignment Scores:
Pred. No.: 1.81e+03 Length: 1662
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-815-242-8100 (1-1662)

Oy 1 GlyTyr**ValGluGlu 6
      ||||| ||||| ||||| |||||
Db 808 GGATACCTCTGTAGAAGAA 825

RESULT 63
US-10-282-122A-7877
; Sequence 7877, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7877
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-282-122A-7877
```

Alignment Scores:
Pred. No.: 1.81e+03 Length: 1662
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-282-122A-7877 (1-1662)

QY 1 GlyTyr***ValGluGlu 6
DB 808 GGATACCTCTGTAGAGAA 825

RESULT 64

US-10-437-963-70740
; Sequence 70740, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 70740
; LENGTH: 1728
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71282C.1
US-10-437-963-70740

Alignment Scores:
Pred. No.: 1.89e+03 Length: 1728
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-437-963-70740 (1-1728)

QY 1 GlyTyr***ValGluGlu 6
DB 1495 GGGTACACGGTGGAGGAG 1512

RESULT 65

US-10-424-599-24574/c
; Sequence 24574, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 24574
; LENGTH: 1781

; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_122194C.1
US-10-424-599-24574

Alignment Scores:
Pred. No.: 1.95e+03 Length: 1781
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-424-599-24574 (1-1781)

QY 1 GlyTyr***ValGluGlu 6
DB 973 GGGTATACGGTGAAGAG 956

RESULT 66

US-10-424-599-59017/c
; Sequence 59017, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 59017
; LENGTH: 1796
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_24303C.1
US-10-424-599-59017

Alignment Scores:
Pred. No.: 1.97e+03 Length: 1796
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-424-599-59017 (1-1796)

QY 1 GlyTyr***ValGluGlu 6
DB 521 GGGTATGCTGTCGAGGAG 504

RESULT 67

US-10-425-114-29590
; Sequence 29590, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28

```
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 29590
; LENGTH: 1949
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY077B08_FLI
US-10-425-114-29590

Alignment Scores:
Pred. No.: 2.15e+03 Length: 1949
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-425-114-29590 (1-1949)

Qy 1 GlyTyr***ValGluGlu 6
Db 1742 GGGTACTCCGTGGAAGAG 1759

RESULT 68
US-10-027-632-99250/c
; Sequence 99250, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99250
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-99250

Alignment Scores:
Pred. No.: 2.19e+03 Length: 1976
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 13 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-027-632-99250 (1-1976)

Qy 1 GlyTyr***ValGluGlu 6
Db 1558 GGATACCTCAGTTGAGGAA 1541

RESULT 69
US-10-027-632-99250/c
; Sequence 99250, Application US/10027632
```

```
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99250
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-99250

Alignment Scores:
Pred. No.: 2.19e+03 Length: 1976
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-5 (1-6) x US-10-027-632-99250 (1-1976)

Qy 1 GlyTyr***ValGluGlu 6
Db 1558 GGATACCTCAGTTGAGGAA 1541

RESULT 70
US-10-424-599-77544
; Sequence 77544, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 77544
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_41039C.1
US-10-424-599-77544

Alignment Scores:
Pred. No.: 2.24e+03 Length: 2019
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservatives: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
```

DB: 16 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-424-599-77544 (1-2019)
Qy 1 GlyTyr***ValGluGlu 6
Db 1742 GGTACTCCGTGGAGAG 1759
RESULT 71
US-10-108-260A-2081/c
; Sequence 2081, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108.260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2081
; LENGTH: 2104
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-2081
Alignment Scores:
Pred. No.: 2.34e+03 Length: 2104
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 16 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-108-260A-2081 (1-2104)
Qy 1 GlyTyr***ValGluGlu 6
Db 2058 GGATATGCAGTAGAGAA 2041
RESULT 72
US-10-369-493-33444
; Sequence 33444, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33444
; LENGTH: 2151
; TYPE: DNA
; ORGANISM: Desulfitobacterium hafniense
US-10-369-493-33444
Alignment Scores:
Pred. No.: 2.4e+03 Length: 2151
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-369-493-33444 (1-2151)

Qy 1 GlyTyr***ValGluGlu 6
Db 1564 GGCTATACCGTAGAGAA 1581
RESULT 73
US-10-369-493-42642
; Sequence 42642, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 42642
; LENGTH: 2226
; TYPE: DNA
; ORGANISM: Anabaena PCC7120
US-10-369-493-42642
Alignment Scores:
Pred. No.: 2.49e+03 Length: 2226
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 15 Gaps: 0
US-10-030-194A-5 (1-6) x US-10-369-493-42642 (1-2226)
Qy 1 GlyTyr***ValGluGlu 6
Db 199 GGTACACGGTCGAGAG 216
RESULT 74
US-09-391-340-9
; Sequence 9, Application US/09391340A
; Patent No. US20020013455A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Mather, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/09/391,340A
; CURRENT FILING DATE: 1999-09-07
; EARLIER APPLICATION NUMBER: US 08/907,166
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Desulfurococcus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2286)
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1801)..(1801)
; OTHER INFORMATION: s at position 1801 is either c or g
US-09-391-340-9
Alignment Scores:
Pred. No.: 2.56e+03 Length: 2289

Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-391-340-9 (1-2289)

Qy 1 GlyTyr***ValGluGlu 6
Db 2065 GGTACACCGTGGAGGAG 2082

RESULT 75

US-09-948-369-9
; Sequence 9, Application US/09948369
; Patent No. US2002012243A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: CALLEN, Walter
; APPLICANT: MATHUR, Eric
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS C
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DIVER1350-3
; CURRENT APPLICATION NUMBER: US/09/948,369
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US 09/656,309
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/391,340
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/907,166
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Desulfurococcus sp.
US-09-948-369-9

Alignment Scores:
Pred. No.: 2.56e+03 Length: 2289
Score: 27.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.43% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-5 (1-6) x US-09-948-369-9 (1-2289)

Qy 1 GlyTyr***ValGluGlu 6
Db 2065 GGTACACCGTGGAGGAG 2082

Search completed: November 3, 2004, 19:56:58
Job time : 296 secs